

THE BURDEN OF GENETIC DISEASE AMONG
HOSPITALIZED CHILDREN IN NEWFOUNDLAND

BRIDGET A. FERNANDEZ

**THE BURDEN OF GENETIC DISEASE AMONG HOSPITALIZED
CHILDREN IN NEWFOUNDLAND**

By

Bridget A. Fernandez

**A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Master of Science in Clinical Epidemiology**

**Faculty of Medicine
Memorial University**

September 29, 2009

St. John's

Newfoundland and Labrador

Thesis Abstract

Background: The Newfoundland and Labrador (NL) population is enriched for certain genetic diseases due to genetic drift and founder effects. Five previous studies have examined the amount of genetic disease among children admitted to a single pediatric hospital. The frequency of strongly genetic disorders (chromosomal and Mendelian disorders) ranged from 4 to ~11%, and all the hospitals served admixed populations. We conducted a similar study, hypothesizing that the amount of genetic disease among hospitalized NL children might be higher than in the previously published studies.

Objectives: We determined the genetic content of 4,144 consecutive hospitalizations to Newfoundland and Labrador's only pediatric hospital. By reviewing the discharge summary, each admission was retrospectively classified into one of 11 genetic content groups. We also compared the utilization of hospital resources by children with strongly genetic versus minimally genetic conditions. Finally, we determined the appropriateness of referrals for genetic services.

Results: Out of 4,144 children, 8.3% had a strongly genetic disease (342 patients with a Mendelian or chromosomal syndrome). Another ~25% (1,033 patients) had a moderately genetic disease and 67% (2,769 cases) were classified as minimally genetic. Children in the strongly genetic group had a mean length of stay (8.01d) that was significantly longer than the non-genetic group (3.99 d), with more cumulative surgeries and cumulative hospital days. Children with single-gene disorders were at risk for prolonged lengths of stay (> 7days). Of 3,281 unique admissions, 1 in 4 children with a diagnosis that is an indication for genetic consultation failed to be referred. The largest deficit occurred for

children with mental retardation and for those with birth defects that have a significant genetic component.

Conclusions: The proportion of admitted children with chromosomal and Mendelian disorders was comparable to previous studies, so that even though certain genetic diseases are overrepresented in NL, this is not reflected among hospitalized children. Our dataset may be enriched for children with high heritability multifactorial diseases (including diabetes and asthma), but comparison with earlier studies is complicated by differences in classification schemes and/or by the fact that the incidence of some of these diseases has increased over the past three decades.

Acknowledgements

Dedication:

I would like to dedicate this thesis to my parents (Peter and Dzintra Fernandez) who have always been encouraging of my work, and to my husband (Michael Woods) for his patience with me and with this project.

Acknowledgements:

I am grateful to my supervisors, Drs. Brendan Barrettp, Patrick Parfrey and Proton Rahman for their guidance and mentorship. I have learned a great deal about epidemiology from each of them.

I would also like to thank Ms. Susan Stuckless (statistician) who in spite of a heavy workload has always made time to help me with the statistical analysis for this and several other research projects.

I would like to extend special thanks to Ms. Kathy Whitten (research nurse) for meticulously reviewing and coding over 4,000 discharge summaries and for her patience with me. This was the first study I designed and there were several issues with the methods and the database that were altered part way through the study, so that Kathy had recode a number of cases.

I would like to thank other members of the research team including Ms. Barbara Noble who enrolled 200 children and their parents into the validation study and kept this part of

the project on track. Ms. Andrea Kavanagh was always able to quickly rectify glitches with the database and this was also very much appreciated.

Finally, I would like to thank the Hospital for Sick Children's Research Foundation for supporting this project.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
Table of Contents	vi
List of Tables.....	x
List of Figures.....	xii
List of Appendices.....	xiii
Thesis Overview.....	xv
Co-authorship Statement.....	xvii
Chapter 1: INTRODUCTION.....	1
1A Background of study.....	2
1B Hypothesis.....	3
1C Research Questions.....	3
1D Relevance and benefits of the study.....	3
Chapter 2: LITERATURE REVIEW.....	5
2A The genetic basis of human disease.....	6
2A.1 Recurrence risk counseling.....	8
2A.2 Estimates of the degree of underlying genetic susceptibility for multifactorial diseases: Lambda s and Heritability.....	10
2B The Newfoundland disease heritage.....	11
2B.1 Persistent geographic isolation and homogeneity leading to high kinship coefficient in particular rural communities.....	12

2B.2.1 Genetic drift producing high frequencies of certain deleterious disease alleles in the Newfoundland population – Mendelian disorders.....	13
2B.2.2 Genetic drift producing high frequencies of certain deleterious disease alleles in the Newfoundland population – Multifactorial diseases.....	16
2B.3 The presence of founder effect for certain Mendelian disorders and probably for some complex traits.....	19
2C How does the Newfoundland population compare with other founder populations with respect to age?.....	22
2D Previous studies of the burden of genetic disease among hospitalized children.....	25
2D.1 Five previous studies of the burden of genetic disease among children admitted to a single tertiary-level pediatric hospital.....	27
2D.2 Conclusions from previous pediatric inpatient studies.....	34
Chapter 3: METHODS & ANALYSIS.....	39
3A Methods -Validation study.....	40
3B Methods – Main study.....	41
3B.1 Study population and study subjects.....	41
3B.2 Sample size calculation.....	42
3B.3 Data extraction and outcomes.....	44
3C Analysis.....	46
3C.1 Analysis of the validation study data.....	47
3C.2 Determination of the genetic content of 4,144 hospitalizations.....	47
3C.3 Analysis of hospital utilization data.....	48

3C.4 Multivariate regression to test the hypothesis that the genetic content of an admission influences the length of stay.....	49
3C.5 Analysis of referral pattern data.....	51
Chapter 4: RESULTS.....	52
4A Results of validation study.....	53
4B Determination of genetic content of 4,144 separations.....	55
4B.1 Description of the distribution of the genetic content of 4,144 consecutive separations and 3,281 unique separations.....	55
4B.2 Description of the most frequent specific diagnoses associated with 4 of the 11 hierarchical final categories.....	62
4C The impact of an admission's genetic content on the utilization of hospital resources.....	66
4D Multivariate regression to test the hypothesis that the genetic content of an admission influences length of stay.....	69
4E Clinical genetic referral rates among hospitalized children.....	73
Chapter 5: DISCUSSION.....	75
5A Validation study.....	76
5B The genetic content of 4,144 consecutive separations.....	78
5B.1 Comparison of the genetic content of the consecutive dataset to previous studies.....	80
5B.1.1 Limitations.....	81
5B.1.2 Discussion.....	91

5B.1.3 Comparison of the eight genetic content categories from this study's consecutive dataset to Hall et al. 1978.....	96
5B.2 The most frequent Mendelian and chromosomal diagnoses for 3,281 unique separations.....	97
5C Impact of the genetic content of an admission on the utilization of hospital resources.....	101
5D ANCOVA regression model.....	104
5E Referral rates for clinical genetics services among hospitalized children.....	106
5E.1 Strongly genetic diagnoses (IA, IB, ID, IE).....	106
5E.2 IC diagnoses (teratogenic exposures).....	110
5E.3 IIA diagnoses (high heritability birth defects).....	110
5E.4 IIIA diagnoses.....	117
5E4.1 IIIA diagnoses – autism.....	118
5E4.2 IIIA diagnoses – mental retardation.....	120
5E.4.3 Other IIIA diagnoses.....	121
5E.5 Deficits in genetic consultations versus existing clinical resources.....	121
5E.6 Future of genetic testing for complex diseases and the role of medical geneticists/genetic counselors.....	124
5E.7 Pharmacogenomics.....	130
5F Strengths and limitations of the study.....	131
5G Conclusions.....	135
References.....	141

LIST OF TABLES

Table 2-1: Comparison of proportion of genetic admissions between single hospital-based studies.....	35
Table 4-1: Comparison of the categorization of 201 pediatric hospitalizations by discharge summary method compared with gold standard (review by a medical geneticist).....	54
Table 4-2: Number of separations that were categorized differently by the discharge summary method compared with the gold standard, but for which the miscategorization did not change the genetic content group.....	55
Table 4-3: Category breakdown of 9,813 diagnoses for 4,144 consecutive patients.....	56
Table 4-4: 9,812 diagnoses for 4,144 consecutive patients stratified into three groups based on the burden of genetic disease.....	57
Table 4-5: Distribution of “Final Category” for 4,144 consecutive hospitalizations.....	58
Table 4-6: “Final Category” for 4,144 separations, stratified into three groups based on the burden of genetic disease.....	58
Table 4-7: Distribution of “Hierarchical Final Category” for 4,144 consecutive admissions.....	60
Table 4-8: Distribution of “Hierarchical Final Category” for 3,281 unique admissions.....	60
Table 4-9: “Hierarchical Final Category” for 4,144 consecutive separations, stratified into three groups based on the burden of genetic disease.....	61

Table 4-10: “Hierarchical Final Category” for 3,281 <i>unique</i> separations, stratified into three groups based on the burden of genetic disease.....	61
Table 4-11: Description of the most frequent specific diagnoses associated with the “hierarchical final categories” IA, IB, IIA, and IIIA, for 3,281 unique separations.....	63
Table 4-12: Distribution of the modes of inheritance for the specific single gene diagnoses within the IB “hierarchical final category”, representing 127 unique hospitalizations.....	65
Table 4-13: Impact of the “hierarchical final category” on length of stay (LOS) for 4,144 consecutive separations.....	66
Table 4-14: Impact of “final category” on cumulative surgeries and cumulative hospital days, for 4,144 consecutive separations.....	67
Table 4-15: Mean age and length of stay for 4,144 consecutive admissions by “hierarchical final category”	68
Table 4-16: Clinical genetic referral rates for 3,281 unique separations analyzed according to “hierarchical final category”	74
Table 5-1: Comparison of final category group distributions between validation and main study datasets.....	77
Table 5-2: The distribution across seven genetic content categories for six pediatric hospitalization studies including the present study.....	87
Table 5-3: Comparison of frequencies across seven genetic content groups between Hall et al. 1978 and the present study – Chi-squared analysis.....	89

Table 5-4: Comparison of six pediatric separation studies: type of hospital; sample size; method of categorization; and proportion of strongly genetic (group 1) and moderately genetic (group 2) hospitalizations.....90

Table 5-5: Children with a final genetic category that was an indication for genetics consultation who were not referred for genetic services.....123

Table 5-6: Breakdown of patients from the unique admissions dataset (3,281 patients) who will require genetic services once complex disease genetic testing is routinely available.....129

LIST OF FIGURES

Figure 4-1: Proportion of 4,144 consecutive separations classified by ‘Final Category’ into three broader genetic content groups.....59

Figure 4-2: Test of the “homogeneity of regression slopes assumption” showing that there is a significant interaction between two of the model’s variables: “age at admission” and “final genetic content category”.....70

Figure 5-1: Break down of 4,144 consecutive admissions by final genetic content category.....79

LIST OF APPENDICES

Appendix A: Diseases within the 11 genetic content categories (IA-IV).....	A-1
Appendix B: ANCOVA - Case Mixed group (CMG) codes.....	B-1
Appendix C-1: Reanalysis of Day and Holmes (1973) data.....	C-1
Appendix C-2: Reanalysis of Hall et al. 1978 data	C-2
Appendix C-3: Comparison of Categories from McCandless et al. 2004 to present study.....	C-6

LIST OF TABLES IN APPENDICES A-C

Table A-1: Numbers of patients from unique admission dataset of 3,281 patients belonging to each of the 20 most common Case Mixed Group (CMG) codes.....	B-1
Table C-1: Number and diagnoses of patients from Day and Holmes 1973 study belonging to our genetic content categories IIA, IIB, IIIA and IIIB.....	C-1
Table C-2: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIA (high heritability birth defects).....	C-2
Table C-3: Number and diagnoses of patients from Hall et al., 1978 belonging to our genetic content category IIB (low heritability birth defects)	C-3
Table C-4: Number and diagnoses of patients from Hall et al., 1978 belonging to our genetic content category IIIA (high heritability multifactorial diseases excluding birth defects).....	C-4

Table C-5 Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIIB (low heritability multifactorial diseases excluding birth defects).....	C-4
Table C-6: Number and diagnoses of patients from Hall et al 1978 belonging to our genetic content category IV (acquired disease with low or no genetic contribution).....	C-5
Table C-7: Summary of the number of patients from Hall et al, 1978 belonging to our eleven genetic content categories.....	C-5
Table C-8: Diagnoses of patients from McCandless category IC belonging to our genetic content categories IIA, IIB, IIIA, IIIB, ID and IE.....	C-6
Table C-9: Diagnoses of patients from McCandless category IIA belonging to our genetic content categories IIB, IIA and ID.....	C-7
Table C-10: Diagnoses of patients from McCandless category III belonging to our genetic content categories IIIA and IIIB.....	C-7

Thesis Overview

This thesis is presented in five chapters.

Chapter 1 gives a brief description of the rationale for the study, followed by the study's three research questions and our initial hypotheses.

Chapter 2 contains the literature review and is divided into three sections. It begins with an overview of the genetic basis of human disease. Human diseases can be divided in two broad categories: diseases that are completely or nearly completely genetically determined versus multifactorial diseases. The former category includes chromosomal and Mendelian syndromes, along with the more recently delineated genomic disorders. The implications for recurrence risk counseling are also reviewed. The second section describes Newfoundland and Labrador's unique population structure from a genetic perspective, emphasizing that it is one of the few relatively young founder populations that has been recognized in the literature. Founder effect has led to over-representation of several Mendelian syndromes which are summarized. The third section reviews the five previous single-hospital based studies (published from 1973 to 2004) which examined the amount and burden of genetic disease among hospitalized children.

Chapter 3 contains a description of the study methods. Prior to beginning the main project, a validation study was performed. The genetic content of 201 randomly selected pediatric hospitalizations from the year 2000 were independently categorized in two ways: firstly using the research nurse/discharge summary method (which was later also used in the main study) and secondly by a gold standard method, i.e. review of the chart

and patient by a medical geneticist. Then the main study methods are described including: the study population; the sample size calculation; and the data extraction and outcomes. Chapter 3 also presents the methods used for analyzing the data. The analysis methods are divided into five sections: analysis of the validation dataset; determination of the genetic content of the main dataset (4,144 consecutive hospitalizations to the Janeway Hospital from an 14-month period, representing 3,281 unique separations); analysis of hospital utilization data; regression analysis to test the hypothesis that the genetic content of an admission influences length of stay; and analysis of the genetics referral pattern data.

The results are presented in **Chapter 4**, following the format laid out in Chapter 3.

Finally, **Chapter 5** contains a discussion of the above results, beginning with the validation study followed by those of the main study. Where possible, our results were compared with previous pediatric hospitalization studies, all of which used samples drawn from admixed populations. This comparison includes reanalyses of data from three of earlier studies (Day and Holmes 1973; Hall et al. 1978; McCandless et al. 2004). Some of the commonest strongly and moderately genetic diagnoses in our unique admission dataset are considered in the context of this province's population structure. Then, we discuss hospital utilization data. Because of the strength of our length of stay (LOS) data, we used multivariate regression to examine the impact of the genetic content of an admission on LOS. The referral rates for medical genetic consultation and referral deficit patterns are considered. Finally the strengths and weaknesses of this study are presented.

Co-authorship statement

- 1) *Design and identification of the research proposal:* The research topic was identified by Dr. Bridget Fernandez who was also primarily responsible for the study design. Along with Dr. Proton Rahman (Professor), she was co-principal investigator of the grant that was used to fund the study. Modifications to the study design and analysis were suggested by Drs. Proton Rahman (Professor), Brendan Barret (Professor) and Patrick Parfrey (Professor) and these were incorporated into the study.
- 2) *Practical aspects of the research:* The validation study involved prospective review of 201 pediatric admissions and the gold standard assessment (examination of the patient and review of the medical chart) was completed by Dr. Bridget Fernandez with review of the discharge summaries by Ms. Kathy Whitten (research nurse). The patients for the validation study were enrolled by Ms. Barbara Noble (research nurse). For the main study, Ms. Kathy Whitten reviewed and categorized 4,144 hospitalizations and each categorization was then reviewed by Dr. Bridget Fernandez. An access database for the study was designed by Mr. Craig Harding. Additional database support was provided by Ms. Andrea Kavanagh. All data entry for the study was completed by Ms. Kathy Whitten.
- 3) *Data analysis:* Ms. Susan Stuckless (statistician) and Dr. Bridget Fernandez completed the data analysis and statistical analysis for this study, with input from Drs. Brendan Barrett (Professor), Proton Rahman (Professor) and Patrick Parfrey (Professor).

- 4) *Manuscript preparation*: Dr. Bridget Fernandez performed the literature review and wrote the thesis with revisions made by Drs. Brendan Barrett (Professor), Proton Rahman (Professor) and Patrick Parfrey (Professor).

CHAPTER 1:

INTRODUCTION

1A. Background of study: Newfoundland and Labrador (NL) is recognized as having a population that is enriched for certain genetic disorders. The province has a population of ~509,000, and already has 27 entries in Online Mendelian Inheritance of Man (www.ncbi.nlm.nih.gov/omim/) as examples of disorders whose genetic basis was significantly elucidated using Newfoundland families.¹ The number of entries is 10-fold higher than that of the neighboring three provinces (Nova Scotia, New Brunswick and Prince Edward Island) when adjusted for population size.

Previous studies from Mexico and the United States revealed that genetic disease is a significant cause of admission to pediatric hospitals (Day et al. 1973; Scriver et al. 1973; Hall et al. 1978; Carnevale et al. 1985; McCandless et al. 2004). The economic burden of patients with genetic disease was also documented. However, these studies were performed in large tertiary (and in some cases quaternary) hospitals serving admixed metropolitan populations. No such data exists for Newfoundland. Obtaining this data has immediate implications for managing the health of NL children and for planning the resources needed for health care delivery. Because hospitalization is one of the major sequelae of morbidity, we examined over 4,000 consecutive pediatric separations from the only tertiary children's hospital in the province (the Janeway Children's Health and Rehabilitation Centre).

¹ In the genetic literature, our province's population is commonly referred to as "the Newfoundland population" (for examples see Rahman et al. 2003¹; Mannion 1977; Martin et al. 2000; Moore et al. 2008). Through this thesis, the term "Newfoundland population" or the abbreviation "NL" is used.

1B: Hypothesis: Newfoundland has a unique population structure with recognized founder effects. We hypothesized that the amount of genetic disease among hospitalized children in this province might be higher than in hospitalized children from admixed populations. Furthermore, we suspected that children admitted with genetic diseases are under-referred for genetic services and that they collectively utilize more hospital resources than children with non-genetic conditions. To answer these questions we reviewed and described the genetic content of 4,144 consecutive admissions² to Newfoundland's only tertiary pediatric hospital, the Janeway Children's Health and Rehabilitation Centre.

1C Research Questions:

1. What is the burden of genetic disease among hospitalized children in Newfoundland?
2. Do pediatric patients with genetic diseases utilize more hospital resources than children with non-genetic conditions?
3. What is the capture/ referral rate to the provincial medical genetics service for patients with strongly genetic diseases?

1D Relevance and benefits of the study: Our findings will generate a number of immediate benefits. We have highlighted the burden that genetic diseases impose on this province's pediatric health care system, both in terms of inpatient care and medical genetics resources. This information can be used by government officials and hospital

² Through this thesis, the terms "admissions to hospital", "separations from hospital" and "hospitalizations" are used interchangeably.

administrators as they make decisions about allocation of funds and recruiting pediatric and medical genetic health care professionals. In addition, we used regression analysis to test the hypothesis that the genetic content of a pediatric hospitalization impacts length of stay. While we found that the specific admitting diagnosis explained the largest proportion of the variance, we determined that children with Mendelian diseases who are admitted for any reason are at high risk for prolonged length of stay (> 7 days), and this information should be useful for hospital discharge planners. Finally, we estimated the proportion of hospitalized children who failed to receive appropriate genetic services (1 in 4). Recognition of referral deficit patterns will allow these to be addressed through education of the referring health care providers.

CHAPTER 2:
LITERATURE REVIEW

2A The genetic basis of human disease: When medical geneticists formulate a differential diagnosis, two broad disease categories are considered: strongly genetic disorders versus weakly genetic/acquired disorders. The strongly genetic group includes Mendelian, chromosomal and mitochondrial diseases plus the more recently delineated category of genomic disorders.

Mendelian diseases are those that are largely determined by the presence of a mutation or mutations in a single gene. These are further classified into one of three inheritance patterns: autosomal dominant; autosomal recessive and X-linked. While we now appreciate that the severity of a Mendelian disease may be determined by one or more additional genes (genetic modifiers), the presence or absence of the disease is determined by the genotype at one particular genetic locus. For example, variants in the modifier gene *CFMI* determine the presence or absence of meconium ileus in infants who carry recessive mutations in the *CFTR* gene and who have cystic fibrosis (Zielenski et al. 1999).

Chromosome disorders also belong to the strongly genetic category. These are disorders where the phenotypes are largely determined by physical changes in either chromosome number (e.g. trisomies, the commonest of which is Down syndrome), chromosome structure (e.g. translocations, inversions) or in chromosome origin i.e. uniparental disomy, which for example explains some cases of Prader Willi syndrome.

Mitochondrial diseases are produced either by mutations in the circular mitochondrial genome (e.g. Leber Hereditary Optic Neuropathy) or by mutations in nuclear DNA genes

that encode proteins that are imported into the mitochondria (e.g. autosomal recessive Complex I deficiency) (Harper 2003¹).

Genomic disorders are a relatively recent addition to the group of strongly genetic conditions. Affected individuals are usually “syndromic” looking and/or developmentally delayed. These diseases are caused by genomic rearrangements (mostly mediated by non-allelic homologous recombination) which causes a segment of the genome to be present in the “wrong dose” (Stankiewicz and Lupski 2003). These microdeletions and microduplications are too small to be identified by routine chromosome analysis. Over the past five years, it has become possible to screen the entire human genome for imbalances that would have escaped detection on routine karyotype using a DNA chip technique called genomic microarray analysis.

The strongly genetic disorders described above have been categorized as “group 1” diagnoses in the study that follows.

Most common diseases do not follow Mendelian patterns of inheritance, but it has long been recognized that many show familial clustering, which in some cases is striking. These non-Mendelian common diseases include birth defects like congenital heart disease and neural tube defects, as well as chronic diseases (examples include asthma, diabetes, hypertension and autism). Various terms are used for these in the literature including “multifactorial”, “polygenic” and “complex” diseases. In his textbook entitled *“Practical Genetic Counseling”*, Peter Harper has written that the most appropriate term for these is “multifactorial” because this term recognizes that these diseases are caused by both genetic and environmental factors. We believe that most of these conditions are caused

by multiple genetic factors ("susceptibility alleles"), each one being of small or moderate effect (Harper 2003²).

Type I diabetes (T1D) is one of the best understood diseases in this group. At least 20 different susceptibility loci have been identified, the two most important of which are *IDDM1* (the HLA locus) and *IDDM2* (the insulin gene). Polymorphisms of two particular HLA genes (DR and DQ) explain 40-50% of this disease's heritability with a "short class" VNTR polymorphism of the insulin gene explaining 10% (Dubois La Forgue et al. 1997; Noble et al. 1996). Sheehy et al. (1989) showed that the HLA DR3/4-DQ2/8 genotype was present in 2.3% of Colorado newborns, but in > 30% of diabetic children. While the population risk of T1D is 1/300, the risk for newborns with this genotype is 1/20. Although babies with this genotype are in an extreme risk group, the majority will not develop T1D and because no preventive environmental strategy has been identified, multifactorial genetic testing for T1D is not being done outside of research protocols.

In the study which follows, the more strongly heritable multifactorial disorders have been classified as either **group IIA** (for birth defects) or **group IIIA** (for non-congenital conditions). See *section 3B.3*

2A.1 Recurrence Risk Counseling: One of the issues that is addressed as part of a medical genetic consultation is the risk that parents of the proband face if they decide to have another child.

Risk figures for Mendelian disorders are derived from a biologic understanding of the genetic etiology of the disease. Parents of a child with an autosomal recessive disorder have a 1 in 4 chance of having another affected children. A mother with an autosomal dominant disease, like Marfan syndrome, has a 50% chance of passing the condition on to a child of either sex and so forth.

At present, “empiric risk” figures are given for most multifactorial conditions, and genetic testing for the presence or absence of particular susceptibility alleles is not used to give a more precise risk figure. The term empiric means that observational data has been collected about reproductive outcomes for the parents of probands with a particular condition. Provided that the study has been carefully designed and is unbiased, these types of “observed” risk figures are used for genetic counseling, and this practice will continue until the genetic basis each of these multifactorial diseases is fully understood. Unlike Mendelian risk figures which can be universally applied, empiric risk figures are population specific and may be altered by factors like disease severity, the number of affected family members or the sex of the affected individual (Harper 2003²).

2A.2 Estimates of the degree of underlying genetic susceptibility for multifactorial diseases: Lambda s and Heritability. Lambda s (λ_s) and heritability (h^2) provide information about whether a disease is likely to have a genetic component, although each has limitations.

Lambda s (λ_s) is defined as the prevalence of disease in siblings of the proband divided by the population prevalence of the disease. It is a measure of familial aggregation which is influenced by both shared genes and shared environment. Nevertheless in his 1990 landmark paper, Neal Risch suggested $\lambda_s > 2$ as a starting point for separating genetic from non-genetic diseases. Later, Lander and Schork (1994) stated that the genetic mapping of a complex trait is more likely to succeed if the disease has a high λ_s (>10) versus a low one (<2).

The concept of heritability was originally developed to quantify the role of genetic differences in determining the variability of quantitative traits, like obesity. In contrast to λ_s which is a measure of familial clustering, heritability is defined as the fraction of the total phenotypic variance which is caused by genetic factors. Operationally, heritability is estimated from twin studies as $2[r_{MZ} - r_{DZ}]$, where r is the correlation of liability derived from concordance rates between monozygotic (MZ) and dizygotic (DZ) twin pairs.

The heritability calculation assumes that the environmental variation between identical (MZ) and fraternal (DZ) twins is equal. In reality, there is often less environmental variation between MZ than DZ pairs so that the calculation may overestimate the effect of heritable factors (Jarvinen and Aho 1994; Hawkes 1997).

Our present understanding of the genetics of type I diabetes was summarized above (*section 2A*). This disease has heritability estimates of 0.72-0.74 with a λ_s of 15 (Kaprio et al. 1992; Kyvik et al. 1995).

For this study, multifactorial diseases were somewhat arbitrarily divided into two groups based on the degree to which the disease is estimated to be genetically determined. The more strongly genetic group contained diseases that have either a heritability estimate $\geq 50\%$ or a $\lambda_s \geq 10$. See *section 3B.3*

2B The Newfoundland Disease Heritage: The Newfoundland population is genetically enriched. It is primarily composed of descendants from a small number of Irish and Southwest English immigrants who settled in the “New World” around the mid-1700’s. Starting from about 20,000 settlers in 1760, the population grew by natural expansion to ~200,000 in 1890 (Mannion 1977). The current provincial population is approximately 509,000 with 95% of English or Irish descent (Statistics Canada, 2008). This rate of expansion is comparable to that of other “young” genetic isolates including the Central Valley Costa Ricans and the Afrikaners of South Africa (Escamilla et al. 1996; Rahman et al. 2003¹). Approximately 60% of the population of Newfoundland resides in communities with 2,500 inhabitants or less (Statistics Canada, 2001).

The following population characteristics have contributed to a high prevalence of certain genetic disorders in Newfoundland:

2B.1 Persistent geographic isolation and homogeneity leading to a high kinship

coefficient in particular rural communities: The cod fishery was responsible for the settlement of Newfoundland, with most immigrants arriving between the mid-1700's and the early 1800's. They came from two main areas: southeast Ireland and southwest England. By the 1830's the major migrations had concluded, and the population of Newfoundland was ~ 75,000. From this point onwards, the population grew mainly by natural expansion. Several factors contributed to keeping related families together including: geographic isolation due to lack of roads connecting costal communities; religious segregation; and limited immigration (Mannion 1977; Citizens 1977).

A study of the degree of relatedness in three representative outports (two on the island part of the province and one in Labrador) was published by Bear et al. in 1987. The authors found that the overwhelming majority of parents originated from the community in which they were raising a family. For example, in the eastern island outport only 7-9% of the parents came from outside the study area, as did only 0.03% for the west coast outport.

Bear used the same east and west coast outports to show persistent homogeneity of the NL population (Bear et al. 1988). Reconstructed pedigree data was used to calculate the homogeneity coefficient of every individual born in these outports. For individuals born between 1960-79, the average coefficient was 0.0032 and 0.0171 from the east and west outports respectively. These coefficients are consistent with genetic isolation and similar to the coefficients seen in the Hutterites and Amish (Bowen, 1985).

Since the 1980's and the demise of Newfoundland's inshore fishery, there has been significant out-migration from certain outports in the province, with almost no in-migration into any of these communities. However since Bear's studies, geographic isolation has been reduced by modernizing transportation routes. Martin et al. (2000) conducted a more recent study; this group determined the allele frequencies for 12 red cell antigens from individuals from ten outports, and compared these with the parental English and Irish populations. They confirmed persistent genetic isolation in selected NL outports, but also identified heterogeneity in the NL population as a whole.

2B.2.1 Genetic drift producing high frequencies of certain deleterious disease alleles in the NL population – Mendelian disorders: The phenomenon of genetic drift refers to the fluctuation in allele frequency due to chance operating in a small gene pool contained within a small population. Because of the population's size, random factors (e.g. survival, increased fertility) may cause the allele frequency to rise for reasons that have nothing to do with the mutation itself; in larger populations such random effects have a tendency to average out (Nussbaum et al. 2001¹). Because of drift, founder populations often have an increased incidence of particular rare (often autosomal recessive) diseases. This is the case in Newfoundland for a number of usually rare autosomal and X-linked Mendelian disorders. Up until the present generation, high sibship sizes also contributed to increasing the frequency of particular autosomal dominant diseases.

Examples of over-represented single-gene conditions in Newfoundland include: Bardet-Biedl syndrome (Moore et al. 2005); infantile neuronal ceroid lipofucinosi (Moore et al. 2008); I-cell disease (Provincial Medical Genetics Program, *Eastern Health* - PMGP database³); Krabbe leukodystrophy (PMGP database); type 1 multiple endocrine neoplasia (Olufemi et al. 1998; PMGP database) and arrhythmogenic right ventricular dysplasia (Merner et al. 2008; PMGP database).

The example of Bardet-Biedl syndrome is discussed in more detail below.

Bardet-Biedl syndrome (BBS): This autosomal recessive disease is a syndromic form of retinal dystrophy. Its prevalence in NL is 1:18,000, compared with 1:160,000 in more admixed Caucasian populations of northern European ancestry (Moore et al. 2005; Woods et al. 1999; Young et al. 1999). For a monogenic disease, BBS is unusually genetically heterogeneous, with 12 known genes, *BBS1-BBS12* (Ross and Beales 2007).

Over a 20-year period, a team of Newfoundland investigators assembled a local cohort of BBS patients and followed them with serial phenotypic measures (Harnett et al. 1988; Green et al. 1989; O'Dea et al. 1996; Moore et al. 2005). Given Newfoundland's population structure, it would have been reasonable to hypothesize that all 46 BBS patients in the cohort would link to the same BBS locus, and even that most affected individuals would have the same homozygous mutation. An overview of the genetic studies on these 46 patients was published by Moore et al. in 2005, and there was an

³ The Provincial Medical Genetics Program (PMGP) database was created when the provincial genetics service was established in 1986. It captures demographic information from every individual who has been referred for genetic consultation, including those who are waiting for assessment and those who have declined appointments. It contains fields for referring diagnosis, final diagnoses and genetic test results.

unexpected amount of genetic heterogeneity. DNA was available from 40/46 affected individuals and at least *six* BBS loci were implicated. The mutations that were identified included: one mutation in the *BBS1* gene; two mutations in *BBS2*; and four mutations in *BBS6*. Another ten patients came from families that linked to either *BBS3* or *BBS5* (at this stage the genes at these loci had not been identified) and four patients from three families had been excluded from all the BBS loci that were known at the time (8 loci). Amidst all this heterogeneity, there was some degree of founder effect. Eight BBS patients from six families had the same homozygous *BBS1* mutation (p.M390R) which is a common mutation in the European parental population. Another 12 patients had homozygous copies of one of two *BBS6* mutations; one of these two mutations was present in another four compound heterozygous patients.

The molecular genetics of BBS in Newfoundland taught us an important lesson about studying monogenic disorders in this province, i.e. that in spite of the existence of a series of genetic isolates which are susceptible to genetic drift and founder effects, one cannot assume that any given Mendelian disorder will be genetically homogeneous. In retrospect, a clue to BBS's heterogeneity came from the fact that the cases were not localized to a series of neighboring Newfoundland communities. Rather affected individuals originated from several geographically separate areas on the island portion of the province. One is therefore led to the hypothesis that the province's founders introduced at least ten mutant BBS alleles into the population. The frequency of these alleles may then have been increased by heterozygote advantage and/or by genetic drift, producing a dramatically increased incidence of this usually rare autosomal recessive disorder.

Unexpected genetic heterogeneity has proven to be the case for another over-represented autosomal recessive disorder, infantile neuronal ceroid lipofuscinosis, where 18 NL probands segregated eight different mutations in four NCL genes: *CLN2*, *CLN3*, *CLN5* and *CLN6* (Moore et al. 2008).

2B.2.2 Genetic drift producing high frequencies of certain deleterious disease alleles in the Newfoundland population – Multifactorial diseases: While the value of isolates (like parts of NL) for identifying genes of major effect is unquestioned, the contribution that these can make towards elucidating the molecular genetics of complex (multifactorial) diseases is controversial (Kruglyak 1999; Laitinen 2002; Rahman et al. 2003¹; Varilo and Peltonen 2004). Extrapolating from the experience with single-gene disorders, it has been hypothesized that the molecular genetics of multifactorial diseases in isolates will be less complex than in admixed populations, with enrichment of some loci and diminution of others. In other words, genetic drift may have led to elevated frequencies of susceptibility alleles for particular complex diseases. In *some* instances, this would be expected to increase the occurrence of that disease in the population. In line with this theory, an increased incidence of certain complex diseases has been noted in Newfoundland, for example:

1) Psoriasis: there is a 2 to 3-fold increased incidence of psoriasis (approximately 5%) compared with most other Caucasian populations (Nall et al. 1999).

2) Juvenile insulin-dependent diabetes: Newhook et al. (2004) determined that the Avalon Peninsula of Newfoundland (which contains ~ 45% of the province's children)

has the highest reported incidence of type I diabetes worldwide. For the 0-14 year age group, over the 4-year period from 1998-2002, the incidence per year was $> 40/100,000$. This is far greater than the incidence in admixed US populations ($7-15/100,000$) (Karvonen et al. 2000), and even exceeds the incidence of juvenile diabetes in other recognized founder populations including Sardinia ($36.8/100,000$) and Finland ($36.5/100,000$).

3) Neural tube defects (NTD's): Spina bifida is the paradigm for a multifactorial birth defect with a known susceptibility allele (*MTHFR*) and a firmly delineated environmental risk factor (maternal folate deficiency). Studies in the early 1990's demonstrated that a woman's risk for having a child with an NTD could be significantly reduced by periconceptional folic acid supplementation (MRC 1991; Czeizel and Dudas 1992). In spite of knowing this, public health campaigns promoting daily use of a folate-containing multivitamin for all women of childbearing age (without mandatory food fortification) did not appreciably decrease the incidence of NTD's (Abramsky et al. 1999; Olney and Mulinare 2002). In contrast, data from countries with mandatory folate food fortification programs have shown a 30-50% decrease in NTD incidence post-fortification (Honein et al. 2001; Persad et al. 2002; De Wals et al. 2003).

A Newfoundland NTD database is housed in the Provincial Medical Genetics Program (PMGP). It is updated annually, and multiple sources of ascertainment are used to identify all affected liveborns, stillborns, and terminations of pregnancy. Over the 1991-1996 period (pre-fortification), the incidence of NTD in Newfoundland was 4.67 per

1000 live births, and fell to 1.01 during the 1998-2002 post-fortification period (House et al. 2006).

The Newfoundland NTD data along with that of six other Canadian provinces was analyzed by De Wals et al. 2007. During the pre-fortification period of Jan 1993-September 1997, the incidence of NTD in all seven Canadian provinces was 1.58 per 1,000 births with the incidence in Newfoundland being 4.46. In the post-fortification period (April 2000-December 2002) the Canadian incidence fell to 0.86 and the incidence in Newfoundland fell to 0.75. The rate drop in Newfoundland (3.8) was the highest of the seven provinces.

Improvement in the folate status of pregnant Newfoundland women post-fortification has been proven (House et al. 2006; Friel et al. 1995). One could hypothesize that the relatively high incidence of NTD's in NL prior to flour fortification was entirely attributable to an environmental factor, because the incidence fell to the Canadian national average post-fortification. Alternatively, the NL population may have a higher frequency of particular NTD susceptibility alleles that interact with maternal folate status to produce a child with spina bifida. The latter hypothesis is supported by the fact in Ireland the incidence of NTD is also unusually high and this population has a high frequency of well documented "folate-sensitive" NTD risk factor, the c.C677T variant in the *MTHFR* (Kirke et al. 2004). Given that 90-95% of the Newfoundland population is of Irish or English ethnicity (Dr. Jane S. Green, personal communication), it is tempting to speculate that this allele is also frequent in NL; to date this allele frequency in our population has not been determined.

2B.3 The presence of founder effect for certain Mendelian disorders and probably for some complex traits: Founder effect is one form of genetic drift. If one of the original founders of a new subpopulation happens to carry a rare allele, that allele will have a much higher frequency than it had in the larger parental population. Founder effect is well illustrated by the Old Order Amish, a religious isolate of European descent that originally settled in Pennsylvania. The Amish have a high incidence of certain rare autosomal recessive conditions, including a form of dwarfism called Ellis-van Creveld syndrome. Other striking examples of founder effect include: variegate porphyria (an autosomal dominant late onset disorder) in the Afrikaner population of South Africa; type I tyrosinemia in the Quebec genetic isolate from the Lac Saint Jean region; and choroideremia (an X-linked degenerative eye disease) in the Finnish population (Nussbaum et al. 2001⁵). By definition if the prevalence of a disease in a genetic isolate is increased because of a founder effect, one or a very limited number of mutations (usually occurring on a shared, conserved haplotype) account for the majority of disease causing mutations in that population, even in apparently unrelated individuals.

Founder effect has been observed in Newfoundland for a number of Mendelian disorders including: hereditary non-polyposis colorectal cancer syndrome (HNPCC) (Frogatt et al. 1999; Green et al. 2002; Stuckless et al. 2007); multiple endocrine neoplasia type 1 (MEN1)-Burin variant (Olufemi et al. 1998); hereditary diffuse gastric cancer syndrome (Kaurah et al. 2007); the Twilligate variant of hemophilia A (Xie et al. 2002); arrhythmogenic right ventricular dysplasia (Merner et al. 2008). Further details are given below:

#1 HNPCC: In 1999, an intron 5 splice site mutation in the *MSH2* gene (A→T at nt943+3) was reported in 10/20 Newfoundland families. In contrast this so called “family C” mutation was present in only 8% of UK HNPCC families. In eight of the Newfoundland families, the mutation occurred on a common haplotype consistent with founder effect (Froggat et al. 1999). This founder *MSH2* mutation has subsequently been identified in a total of 12 NL families (Green et al. 2002; Stuckless et al. 2007) and in a population based CRC cohort of 750 probands, there were 26 Amsterdam 1 criteria families, nine of which segregated the family C mutation (Dr. Michael O. Woods, personal communication).

#2 MEN1Burin: One hundred and thirty eight MEN1 patients have been recorded in the Provincial Medical Genetics Program database, and 125 (90.6%) have a specific nonsense mutation in exon 10 of the *MEN1* gene (*R460X*), which is also known as the MEN1-Burin variant mutation. These Burin variant patients come from six families whose ancestors can all be traced to a resettled community in Fortune Bay, adjacent to Burin peninsula and the mutation has occurred on a common haplotype. Their phenotype is characterized by a relatively high frequency of prolactinomas and a low incidence of gastrinomas (Bear et al. 1985; Olufemi et al. 1998; Hao et al. 2004; Dr. Jane S. Green, personal communication; PMGP database).

#3 Hereditary diffuse gastric cancer syndrome: In 2007, Huntsman et al. published the results of mutation analysis of the *E-cadherin* (*CHD1*) gene on 38 families who met the criteria for hereditary diffuse gastric cancer syndrome. The cohort included four NL families who had the same mutation (2398delC). The 2398delC families all originate

from resettled communities within a 100-mile radius off the south east coast of NL (Marishene Island) and from St. Pierre and Miquelon. The four families contained 29 cases of gastric cancer and 16 cases of lobular breast cancer, and all mutation carriers shared a common haplotype consistent with a founder effect mutation (Kaurah et al. 2007).

#4 Mild hemophilia A (Twillingate variant): In 2002, Xie et al. reported a founder mutation in the *Factor VIII* gene (valine 2016 → alanine) which explains the high prevalence of mild hemophilia in the province.

#5. Arrhythmogenic right ventricular dysplasia (ARVD5): Fifteen Newfoundland ARVD families (containing 257 affected people) had the same disease associated haplotype on chromosome 3p25. In 2008, all patients with this haplotype were shown to have the same mutation in the *TMEM43* gene (p.S358IL) (Merner et al. 2008).

Complex trait genetics is a much younger field than its Mendelian counterpart, so that fewer Newfoundland “founder” susceptibility alleles have been identified. However one example, *CARD15*, is discussed below.

The *capsase recruitment domain* gene (*CARD15*) is a susceptibility gene for Crohn’s disease and psoriatic arthritis (Hugot et al. 2001; Rahman et al. 2003¹; Rahman et al. 2003²). Dr. Rahman’s group demonstrated a founder effect for a particular *CARD15* allele in the Newfoundland population. The p.Arg702Trp variant accounts for 71% of the NL *CARD15* Crohn’s mutations, compared with 35% of European *CARD15* mutations and 40 of Quebec *CARD15* mutations (Rahman et al. 2003¹; Lesage et al. 2002; Vermeire

et al. 2002). The p.Arg702Trp variant also accounts for ~70% of the *CARD15* mutations associated with psoriatic arthritis in NL (Rahman et al. 2003²).

2C How does the Newfoundland population compare with other founder populations with respect to age?

The appeal of isolates for molecular genetics research is related to their genetic homogeneity (with reduced allelic diversity), which is usually accompanied by greater environmental homogeneity than that of admixed populations. However, there are important differences between founder populations. These factors influence the isolate's disease allele diversity, which in turn determines its usefulness for gene hunting projects. Distinguishing parameters include: the number of founders; the time before population expansion began; the age and current size of the population; the amount of in-migration; and the presence and timing of historical bottlenecks (e.g. famines, epidemics of infectious diseases, wars) (Laitinen 2002; Varilo and Peltonen 2004).

The "old" genetic isolates are ~200-400 generations old and are usually demographically stable. Examples include the Basques of southwest Europe and the Saami (Lapps) of Scandinavia. The Saami forager population numbers about 50,000 and has been spread across a large land area because of the Saami's source of livelihood. These older isolates have often lost their distinct spectrum of monogenic disorders and so have been largely ignored by molecular geneticists.

Molecular genetic studies searching for alleles that confer susceptibility to complex traits have primarily targeted "middle-aged" genetic isolates. These are populations that had a

restricted number of founders and that are 100-200 generations old. Examples include Finland (early settlement portion is ~ 2,000 years old) and Iceland (~ 1,000 years old).

Newfoundland belongs to the group of “very young” founder populations, which are 10-20 generations old. Other examples include: the late-settled parts of Finland (~330 years old); the Central Valley region of Costa Rica; the Lac Saint Jean region of Quebec (~250-400 years old); the Amish (~ 250 years old); the Canadian and US Hutterites (~ 130 years old); and the inhabitants of the Pacific island of Tristan da Cunha (~ 100 years old) (Rahman et al. 2003¹; Varilo and Peltonen 2004; Laitinen 2002).

Younger isolates are expected to have a distinct pattern of over-represented monogenic disorders. For example in Finland, founder effect has produced a high incidence of 36 monogenic (mainly autosomal recessive) disorders (Peltonen et al. 2004). However this does not necessarily translate to a net increase in the total frequency of Mendelian disorders in isolates. Certain Mendelian conditions are expected to be under-represented because these mutant alleles were not introduced into the population by its founders. To date, there have been no studies of founder populations which have measured the overall burden of genetic disease. Rather families from these populations have been used to identify genes causing specific monogenic disorders or the population has been studied at a theoretical level to estimate its value for the identification of complex trait susceptibility alleles.

Along a similar vein, younger founder populations are not necessarily predicted to have a higher net genetic burden from complex diseases. However if the prevalence of a particular multifactorial disease is higher than in the general population, this might be

explained by the fact that usually rare, but highly penetrant susceptibility alleles have become more common in the isolate because of genetic drift. In 2005, Oen et al. reported that a complex disease, rheumatoid arthritis (RA), is enriched in the North American Native (NAN) population from Manitoba and Northwest Ontario. This isolate is comprised of Cree and Ojibway Indians and has a 5-fold increased rate of RA compared with North American Caucasian and European populations. In addition to a higher prevalence of RA, the NAN have an earlier age of disease onset, greater disease severity, and higher familial prevalence, all suggestive of increased frequencies of RA susceptibility alleles due to genetic drift and founder effect.

In Newfoundland, we would therefore expect to find an increased incidence of certain Mendelian diseases, with other monogenic disorders that occur less frequently than in admixed populations. One might also expect to find over-representation of a select number of complex diseases (see *section 2B.3* for documented examples).

Fragile X syndrome, due to mutations in the X-linked *FMRI/FRAXA* gene, is an example of a monogenic disorder that is *less* common in Newfoundland than in more admixed populations. In most populations, FRAXA syndrome is the most frequent inherited cause of mental retardation, occurring in 1 in 1,250 males and 1 in 2,500 females. Because of its high prevalence, *FMRI* genetic testing is routinely ordered as part of the evaluation of any child with cognitive delay of unknown etiology. There are at least several thousand developmentally delayed individuals who have been evaluated through the Provincial Medical Genetics Program in St. John's and who have had *FMRI* genetic testing. Only one family with molecularly confirmed FRAXA syndrome has been identified, and that

family is not originally from Newfoundland (Dohey et al. 2008; PMGP database). Interestingly the same paucity of FRAXA patients has been reported in Nova Scotia (Beresford et al. 2000).

2D Previous studies of the burden of genetic disease among hospitalized children:

Dating back at least to the 1950's, investigators have been interested in documenting the frequency of genetic disease in human populations. Such information is essential for planning rational health care strategies. In one of the earliest publications from 1959, Stevenson reported that 26% of all institutional beds in Ireland were occupied by patients with genetic disease.

The strongest population-based study was published by Baird et al. in 1988 using data from the British Columbia Health Surveillance registry. This is an on-going population-based registry with multiple sources of ascertainment. The authors estimated the population burden of genetic disease in over 1 million consecutive live births, and concluded that before age 25 years at least 53/1,000 liveborns can be expected to have a disease with an important genetic component. This total included single-gene disorders (3.6/1,000), chromosomal disorders (1.8/1,000) and multifactorial disorders present by age 25 (46.4/1,000).

Hospitalization is one of the most expensive forms of medical care for genetic disease. There have been five previous studies which estimated the amount and distribution of genetic disease among children admitted to tertiary level pediatric hospitals. These are

discussed below and have been summarized in *table 2-1*. The present study shares many of the design features of these earlier studies.

The other major group of pediatric genetic content studies are retrospective reviews of the causes of death of children admitted to either a pediatric hospital or to a children's intensive care unit. The results of three of these studies are as follows: 34% of 523 deaths in a Utah tertiary level pediatric hospital were due to malformations or genetic disorders (Stevenson and Carey, 2004); 19% of 268 deaths in a pediatric ICU in a teaching hospital in Arkansas were due to a heritable disorder (Cunniff et al. 1995); and 45% of deaths over an 11-year period in a neonatal ICU in Kentucky occurred in patients with a major congenital malformation (Stewart and Hersh 1995).

The above studies share some limitations. They are based on retrospective review of death certificates and/or medical records. Hence the authors relied on the healthcare team that cared for the child (which did not necessarily include a geneticist) to recognize and document the presence of a genetic disorder. Because the boundary between genetic and non-genetic disease has become blurred, classification of whether a particular disease was genetic or not was not always unequivocal. If a genetic disease was present, determination of whether the disease contributed to the child's death was sometimes subjective.

For example Stevenson and Carey (2004) reviewed neonatal and pediatric deaths over a 4-year period (1994-1998) in a university teaching hospital that has a neonatal and pediatric ICU and that is a regional referral center for cardiac surgery. Hence compared to a community pediatric hospital, there was undoubtedly referral bias towards children

with more complex medical problems which includes those with genetic diseases. The classification of the deaths was based on chart summaries prepared by the chief resident for monthly morbidity and mortality rounds, and these summaries were reviewed by the authors. There was no separate category for multifactorial diseases that led to death, and these were included in the non-genetic category.

Cunniff et al. (1995) compared review of the medical record with review of the death certificate alone. When the medical record was reviewed, 51 pediatric ICU deaths were classified as attributable to a heritable disorder. The death certificate listed the heritable disorder in only 21 of these 51 cases.

2D.1 Five previous studies of the burden of genetic disease among children admitted to a single tertiary-level pediatric hospital: PubMed was searched using the terms “pediatric hospital and genetic” and “genetic burden and pediatric”. Five single hospital-based studies were identified.

1) Day and Holmes 1973: These authors published a review of 800 hospital records from the Massachusetts General Hospital, including 200 pediatric inpatients. Primary and secondary diagnoses and family history information were extracted from the hospital chart. The authors used a classification system in which each diagnosis was assigned to one of seven categories based on its genetic content: 1) single-gene; 2) chromosomal; 3) polygenic; 4) probably genetic; 5) developmental; 6) unknown and 7) environmental (which included all malignancies except Wilms tumor). Polygenic disorders were defined as conditions *“considered to be due to additive effects of several minor gene*

abnormalities...with a sibling recurrence risk of at least 5%" (e.g. some isolated major congenital abnormalities, asthma, diabetes, psoriasis). The "probably genetic" conditions were defined as diseases which showed familial clustering and for which a genetic etiology had been postulated (e.g. seizure disorders, migraines and benign familial macrocephaly). Developmental disorders were anomalies that arose from abnormal development, but that were not included in the polygenic group (e.g. undescended testes). The unknown category included conditions for which it was not possible to assign either a genetic or environmental etiology (e.g. cerebral palsy, Wilms tumor and mental retardation). Finally, the environmental group included diagnoses for which the authors felt that influences from the environment were the major cause of disease (infections, accidents, most malignancies).

These authors found that 17% of pediatric inpatients had a primary diagnosis that was of genetic origin (i.e. either a single-gene, chromosomal or a polygenic disease). The largest contribution in the genetic group was made by the polygenic disorders. More specifically 13% of the 200 admitted children had a primary diagnosis of a polygenic disorder, and 28% of these had a positive family history for the same disorder. 4% of the admitted children had a primary single-gene or chromosomal diagnosis.

2) Scriver et al. 1973: In the same year, Dr. Charles Scriver's group published a review of the medical charts of 1,089 children admitted to the Montreal Children's Hospital. The charts were randomly sampled over a 12-month period from 1969-70, during which time the total number of admissions was 12,801. The discharge diagnosis and length of stay were extracted by reading the medical record (note that only the final diagnosis was

extracted). The discharge diagnoses were again divided according to their genetic content, this time into four categories: 1) genetic: single-gene, chromosomal or multifactorial diseases – the latter excluded birth defects; 2) congenital malformations; 3) unknown - probably genetic, but could not be assigned to either of the first two categories; 4) non-genetic.

Scriver's group found that 11% of the discharge diagnoses were "genetic", 18.5% were "congenital malformations", and 6.7% were in the "unknown" category. The authors concluded that at least 30% of admissions to their pediatric hospital were due to the effects of "*abnormal gene-environmental interactions*". During that time frame, less than 300 patients were referred to the genetics service, compared with the 4,000 pediatric patients whom they extrapolate were admitted with "genetic disease". They comment on the tremendous under-referral to their genetic service and on the fact that their genetic clinic would be unable to cope with such a referral volume. The paper includes a list of the 17 most common "genetic" discharge diagnoses (the two most common were atopic hypersensitivity and hemophilia A), and the 14 most common congenital malformations (the two most common were hernia and congenital hip dislocation). Among patients hospitalized for more than 10-days, the congenital anomaly group was over-represented. Moreover, patients with "genetic" illnesses accounted for 70% of the children with multiple admissions.

3) Hall et al. 1978: In one of the largest studies to date, Dr. Judith Hall and colleagues reviewed 4,115 admissions to a 200-bed general pediatric hospital in Seattle. The study population was randomly selected from a total of 8,244 children admitted to this hospital

in 1973. The primary source of information was the discharge sheet which listed the primary admission diagnosis, however in over half of the cases the medical chart was reviewed to establish the cause of the admission. A diagnoses list was assembled for each patient and each diagnosis was put into one of the following five categories: 1) clearly genetic disorders – single-gene or chromosomal disorders; 2) multifactorial/polygenic conditions – this category included conditions for which, at the time, multifactorial recurrence risks were given after specific environmental and single-gene entities had been excluded, for example congenital heart disease, seizures, deafness, mental retardation, cleft palate, diabetes, allergies – note that multifactorial birth defects and multifactorial diseases that are not congenital were analyzed together 3) developmental anomalies – malformations for which the recurrence risk had not yet been established, e.g. esotropia, cryptorchidism, renal abnormalities, multiple congenital anomaly syndromes; 4) familial disorders – conditions for which a familial predisposition appeared to exist, but recurrence risk was thought to be less than for categories 1 and 2, e.g. prematurity, cancer; 5) non-genetic disorders e.g. infections, trauma.

Next, Hall's group assigned an overall category to each admission, which was defined as the lowest category number from the diagnoses list (i.e. the diagnosis with the highest genetic contribution from the list).

Of the 4,115 admissions, 4.5% had clearly genetic disorders, 22.1% had multifactorial/polygenic conditions, 13.6% had developmental abnormalities and 13.2% had familial conditions. 46.6% had non-genetic disorders. Patients with clearly genetic disorders (category 1) had an average of 5.3 admissions, compared to 1.6 previous

hospitalizations for patients with non-genetic conditions. The average length of hospitalization was 3.4 days for the genetic patients and 2.5 days for the non-genetic patients. Only ¼ of the category 1 patients received genetic counseling.

4) Carnevale et al. 1985: These authors reviewed 2,945 admissions to a 350-bed general pediatric hospital in Mexico. This was the first survey carried out in a developing country and included neonatal and pediatric admissions. Diagnoses were extracted from the discharge sheet and by reviewing the medical chart. For each patient, the diagnosis selected for category assignment was the one with the highest genetic designation. The authors used 5 categories: 1) single-gene disorders; 2) chromosomal disorders; 3) diseases of probable but complex genetic cause, including birth defects, multiple malformation syndromes, diabetes, asthma; 4) diseases of unknown cause in which role of genetic factors was not known e.g. cancer, prematurity; 5) non-genetic e.g. infections, trauma, malnutrition. Strongly genetic diseases (categories 1 and 2) accounted for 4.3% of all the admissions. Category 3 (probable but complex genetic cause) represented 33.5%. Therefore, patients with genetic or partly genetic disorders accounted for 37.8% of the patients studied. These genetic patients had longer and more frequent admissions.

Yoon et al. 1997: This is a population-based (rather than single hospital-based) study of pediatric separations. This group reviewed hospital discharge data from over 350,000 patients from California and South Carolina, which captured over 95% of the admissions under age 20 years in these two states during 1991. Each case was classified using ICD-9 codes extracted from Medicare claim forms and from the hospital discharge summary. Each admission was categorized into one of 17 broader ICD categories (e.g. “hereditary

metabolic or endocrine disorders”, “CNS or eye defects”). Nearly 12% of the pediatric hospitalizations in the two states combined were due to birth defects or “genetic disease”; the latter was defined as a Mendelian or chromosomal disorder. This group of children stayed on average three days longer in hospital, and had 4.5 times greater in-hospital mortality than children hospitalized for other reasons. This 12% figure is considerably lower than the 30-40% reported by the previously described single hospital-based studies. This almost certainly reflects the fact that sicker more complex children tend to be referred to pediatric hospitals like the ones used in the above papers. Yoon’s study is population-based and includes pediatric admissions to non-specialty hospitals where the case-complexity would have been lower. Moreover, these authors chose a narrow definition of genetic disease (Mendelian, chromosomal and birth defects) and excluded the complex diseases (e.g. diabetes, asthma, epilepsy) included in the studies by Day and Holmes, Scriver, Carnevale and Hall.

5) McCandless et al. 2004: These authors reviewed 5,747 consecutive admissions representing 4,224 unique patients who were admitted in 1996 to a 244-bed pediatric hospital in Cleveland, Ohio. Each patient’s record was reviewed until a genetic diagnosis was identified or eliminated. The initial review was performed using only electronic discharge diagnoses. If the admission could not be assigned to the category IA (single-gene or chromosomal), the entire written medical chart was manually reviewed; over 5,000 charts were reviewed in this way. The categories used for classifying each separation were modified from those used by Hall et al. 1978 and were:

Category I – Underlying conditions with a strong genetic basis

IA – Single-gene or chromosomal

IB – Multifactorial/polygenic conditions (e.g. congenital heart disease, cleft lip/palate, autism) – includes both birth defects and complex traits with a sibling recurrence risk.

IC – Heterogeneous causes, often genetic (MR, seizures) – i.e. mainly complex traits for which sibling recurrence risk has been identified.

II – Birth defects without known genetic basis

IIA – Malformations of unknown etiology (bladder extrophy)

IIB – Teratogenic disorders (FASD)

III: Acquired disorders with genetic predisposition (asthma, diabetes, cancer) –note that the other four studies all classified cancer in one of the non-genetic categories, so that this study's frequency in this category is comparatively high. See *table 2-1*.

IV: Acquired disorders without genetic determinant (e.g. hypoxic ischemic encephalopathy)

V: No pre-existing chronic medical condition (e.g. infection in previously healthy child, trauma)

There were 1,949 admissions in category I, representing 34% of the total admissions. Within this group 10.8% had single-gene or chromosomal disorders and 14.5% had Category IB multifactorial/polygenic conditions. 37% had an acquired disorder with a genetic predisposition (category III). The mean length of stay was 40% longer for patients with a clearly genetic underlying diagnosis than for non-genetic admissions.

2D.2 Conclusions from previous pediatric inpatient studies: In summary, there have been five previous studies that have attempted to quantify the burden of genetic disease among hospitalized children. In each case, the sample was drawn from a single tertiary-level pediatric hospital; however making direct comparisons between these studies is difficult. Firstly, the publications span almost four decades, over which time we have developed a better understanding of the genetic basis of many disorders. Secondly, no two studies have used exactly the same classification scheme for ranking the genetic content of the admissions. Moreover, information about the specific diagnoses within each genetic content category varies between studies, as does the degree of record review that was performed before assigning the hospitalization to a particular genetic content category.

With the above caveats, *table 2-1* below summarizes the genetic and non-genetic disorder frequencies reported in the five studies that preceded this study (Yoon et al. 1997 was excluded because this was not a single hospital-based study). If “strongly genetic” conditions are defined as Mendelian and chromosomal disorders, these studies have identified frequencies of wholly genetic disorders among hospitalized children ranging from 4-10.8%. “Moderately genetic” diseases (birth defects and complex traits for which sibling recurrence risks are given) occurred with frequencies of 19.5-20.9%. Due to differences in the classification schemes used in the Scriver, Carnevale and McCandless studies, the proportion of moderately genetic separations could not be estimated (see discussion *section 5B.1 & appendix C*).

Table 2-1: Comparison of proportion of genetic admissions between single hospital-based studies (proportions are based the total number of hospitalizations in each dataset).

Study	Hospital	Sample size	Single gene (%)	Chromosomal (%)	High heritability birth defects (%)	Low heritability birth defects (%)	Complex disease with sib RR (%)	Complex disease with low RR (%)	Non-genetic (%)
Day & Holmes 1973*	Massach. General Hospital	200	4.0	0	11	21.5 ^a	8.5	9	46.0 (includes malignancies apart from Wilms tumor)
Scriver et al. 1973	Montreal Children's Hospital	1,089	6.9	0.4	< 18.5 ^b	< 18.5 ^b	3.85 ^c	< 70.4 ^d (assumed to include malignancies)	< 70.4 ^d (assumed to include malignancies)
Hall et al. 1978*	Children's Orthopedic Hospital & Medical Center, Seattle	4,115	4.18	0.63	10.2	9.3	10.7	18.2 (includes malignancies)	46.6
Carnevale et al. 1985	National Institute Peds., Mexico	2,945 includes neonates	5.1 ^e	0.85	< 22.4 ^f	< 22.4 ^f	9.3	16.6 ^g	45.4 ^h
McCandless et al. 2004*	Rainbow Babies & Childrens' Hosp. Cleveland, OH	5,747	9.5	1.29	< 24 ⁱ	< 9.44 ^j	< 59.6 ^k	< 45.1 ^l	27.2

* see appendix C for details of how this study's dataset was reanalyzed using the genetic content categories from the present study.

^a 43 cases including 36 "urinary tract abnormalities", some of which we probably would have classified in our study as IIA (e.g. renal agenesis) or IIIA (e.g. vesicoureteric reflux), but the break down of individual diagnoses was not given.

^b Scriver et al. did not distinguish between high heritability congenital abnormalities (e.g. congenital heart disease) and low heritability birth defects (e.g. tracheoesophageal fistula).

^c In this group of 42 cases, there were 24 cases of "atopic hypersensitivity". The remaining 18 diagnoses were not specified.

^d Scriver et al. did not distinguish between low heritability complex traits and non-genetic diseases

^eIncludes 101 children with an identified single-gene disorder + 50 children with multiple congenital anomalies (our study category IE).

^f Congenital defects were listed as broad categories (e.g. congenital heart malformations, gastrointestinal anomalies), so that this group could not be broken down into high and low heritability birth defects.

^g Includes 250 children with cancers/tumors + 37 with “cerebral palsy and other neural system disorders” + 37 with “renal disorders” + 156 with “miscellaneous” diagnoses.

^h This group consists of 1,341 children where the diagnoses were given as broad categories, e.g. “infectious diseases”, “gastrointestinal”, “nervous system”. We likely would have categorized some of the specific diagnoses within each broad group as IIb and others as IV.

ⁱ Higher heritability birth defects appeared in three McCandless groups (IB, IC, IIA). The total number of separations in these three groups comprises 24% of their dataset, but these three groups also include non-congenital diagnoses and low heritability congenital anomalies.

^j Low heritability birth defects appeared in McCandless categories IC and IIA which together accounted for 9.44% of the dataset, but these two group also included other diagnosis (that we categorized as IIA, IIIA, IIIB, ID, and IE).

^k Higher heritability non-congenital multifactorial diseases appeared in three McCandless groups: IB (14.5% of total separations); IC (8.6% of total separations) and III (36.5 % of total separations). However each of these three groups contained other diagnoses not falling within our IIIA group.

^l Lower heritability non-congenital multifactorial diseases appeared in McCandless categories: IC (8.6% of their total separations) and III (36.5% of their total separations). Both McCandless' groups contained other diagnoses not falling within our IIIB group.

In conclusion, over the past three decades, our understanding of the genetics of common and rare diseases has increased. Treatment options for genetic disease have expanded and some formerly fatal conditions are now compatible with long-term survival. Moreover the province of Newfoundland and Labrador has a unique founder population, and one could hypothesize that it contains more individuals affected with either strongly or moderately genetic diseases.

Because knowledge of the amount of genetic disease that exists among hospitalized children is essential for designing health care strategies for the delivery of genetic service, we conducted a study of the burden of genetic disease among hospitalized children using the only tertiary-care pediatric hospital in this province.

Our study design was similar to that of Hall et al. 1973 and McCandless et al. 2004. For each hospitalization, the minimum review included the ICD-9 discharge diagnosis(es) and the discharge summary dictation. We modified the classification scheme by using lambda s (λ_s) and heritability estimates to divide the birth defects and non-congenital multifactorial diseases into more strongly and less strongly genetic categories.

We used this classification method (which included 11 categories) to describe the genetic content of 4,144 consecutive NL pediatric hospitalizations (3,281 unique separations). We examined hospital utilization data by classifying each separation into one of three broader genetic content groups (strongly, moderately or weakly genetic). We used multivariate regression to examine the impact of the genetic content of an admission on length of stay (LOS). Finally, we described referral patterns among these hospitalized children, identifying groups of children who are under referred for genetic services.

As a prelude to the main part of this project, we conducted a validation study comparing our proposed method of determining the genetic content of a hospitalization to a gold standard (face-to-face interview and record review by a medical geneticist). Medical geneticists are physician specialists who are trained in the recognition and investigation of patients with strongly genetic diseases. We found that our study method had high enough sensitivity and specificity to allow us to apply it to the entire study sample.

CHAPTER 3:
METHODS & ANALYSIS

3A. METHODS - VALIDATION STUDY: Administrative databases are suboptimal for data extraction. Therefore, we embarked on a validation study to examine the accuracy of the retrieval of genetic diagnoses from hospital discharge summaries.

In the main study, review of the index discharge summary by a research nurse was used to identify diseases that have a genetic component. This methodology is limited by the fact that it relies on the admitting pediatrician and/or house staff to recognize the presence of a genetic disorder in the admitted child and to mention this in the discharge summary, regardless of whether this influenced the child's hospitalization (for example polydactyly in a child admitted with gastroenteritis). We propose that a clinical interview of the parent and child by a medical geneticist, coupled with review of the inpatient record by this geneticist forms the best "gold standard" against which to judge the accuracy of other means of identifying such diseases. While this may be the case, it is not feasible to use this method for over 4,000 cases.

We performed a validation study in which a random sample of **201 pediatric cases** (from January – October 2000) was reviewed by a gold standard method and by the research nurse method proposed for the main study.

The participants were enrolled on one day per week over a 10-month period (an average of six children per week). The weekday selected depended on the availability of the geneticist (BF). On the selected day, the research nurse obtained the admission list from the previous night and a random numbers list was used to select the cases.

For each consenting case, the parent completed a face-to-face interview with the medical geneticist prior to discharge from hospital. The geneticist examined the child, reviewed the inpatient record and assigned the patient to one of 11 final genetic categories (described in *section 3B.3*). These cases were independently categorized by the project's research nurse through review of the ICD9 discharge codes, the index discharge summary and up to two previous discharge summaries (i.e. following the protocol outlined in *section 3B.3*). The accuracy of the latter method for "final category" and overall group assignment (group 1, 2 or 3, as defined in *section 3C*) was calculated.

3B. METHODS – MAIN STUDY

3B.1 Study Population and Study Subjects: Pediatric separations from the Janeway Children's Health and Rehabilitation Centre⁴ were reviewed. Located in St. John's, the provincial capital city of Newfoundland and Labrador, this 80-bed facility is the only tertiary-level pediatric hospital in the province. Sixty percent of all pediatric separations in the province occur from this hospital. Moreover, clinical genetic referrals are handled by a single provincial medical genetics program (PMGP), which is based at the Janeway Hospital with two outreach clinics. The program was established in 1986 and data on every referred patient has been entered into a database. Through cross-referencing the patients in the study population with the medical genetics database, those who received genetic services were easily identified. The database also captures children who have been referred to the PMGP who are waiting for an appointment or who failed to keep the appointment that was offered.

⁴ Also referred to as the "Janeway Hospital".

The study population includes all pediatric patients admitted to the Janeway Children's Health and Rehabilitation Centre over a 14-month period between July 2000 to September 2001 (4,144 separations). A pediatric admission was defined as one involving a person between the ages of 1 month and 17 years. Neonates were excluded for several reasons. Firstly, those born in hospital and admitted to the nursery for routine newborn care are a fundamentally different population (that is essentially healthy), compared with children hospitalized because of a medical problem. Neonates admitted to the neonatal ICU represent a group that is potentially enriched for birth defects and other genetic diseases and is a patient population that warrants its own study.

3B.2 Sample size calculation: The study's sample size was calculated to ensure that the smallest category from previous studies (the chromosomal group) would be adequately represented. The percent of children with a chromosomal disorder from four of the earlier studies was 0.4% (Scriver et al), 0.6% (Hall et al. 1978); 0.9% (Carnevale et al. 1985) and 1.29% (McCandless et al. 2004). For the sample size calculation, π was set at 1.3% and the width of the 95% confidence interval was arbitrarily set at 0.4%.

The sample size formula for estimation of a population proportion is:

$$\text{Sample size} = \pi (1 - \pi) (Z_{\alpha/2} / CI)^2$$

Where π = the smallest proportion that should be detected with a high level of confidence. $\pi = 1.3\%$; CI selected at 0.4% and $\alpha = 0.05$

$$\text{Sample size} = 0.013 \times (1 - 0.013) \times (1.96/0.004)^2 = 3,081 \text{ separations}$$

The chromosomal frequency in the above four studies was determined from the total number of separations which included some children who had been hospitalized more than once in the study period. Nevertheless, we interpreted the sample size calculation conservatively and set the sample size at a minimum of ~ 3,100 *unique* admissions.

We reviewed 4,144 consecutive separations. 863 of these were repeat admissions of the same child, leaving 3,281 unique hospitalizations.

3B.3 Data extraction and outcomes: The Newfoundland and Labrador Centre for Health Information Services (NLCHI) reviewed its data tapes from the Janeway Children's Health and Rehabilitation Centre over the study period. Lists of unique separations were provided, including the following discharge data: the ICD9 discharge diagnosis codes, length of stay, number of surgical procedures, total cumulative inpatient days (from birth up to and including the index admission), and the "case mix group" (CMG) code. The CMG code is assigned by the Canadian Institute for Health Information (CIHI), and is described in more detail in *section 3C.4*.

For each separation, the most recent discharge summary was reviewed. If the patient was admitted previously, the research nurse also reviewed the two most recent previous discharge summaries. The attending physician or house staff dictates the discharge summary when the patient is released from hospital. On average, these are 1-3 pages in length and each summarizes the patient's course in hospital including the results of tests and the final diagnoses. Almost all the discharge summaries over the study period were available electronically, so that retrieval of the hospital chart was rarely necessary. Based on the ICD9 discharge codes, review of the index discharge summary and review of up to two previous summaries, a list of diagnoses was determined for each index separation.

For each hospitalization, each diagnosis from the list was assigned to one of the following 11 categories (see *Appendix A* for list of the diseases in each category):

IA: Chromosome disorder ¹

IB: Single-gene disorder ²

IC: Teratogen³ – included within category I because such patients often have birth defects and/or dysmorphic features. Differential includes chromosomal and Mendelian syndromes so that genetic consultation is indicated.

ID: Genetic syndrome identified, but precise mechanism of inheritance unknown² (e.g. Hallerman-Streiff syndrome)

IE: Unidentified genetic syndrome likely exists and referral to a geneticist is indicated (e.g. multiple congenital anomalies; developmentally delayed child who has a birth defect)

IIA: Multifactorial birth defect with known sibling recurrence risk⁴ (e.g. cleft lip).

IIB: Birth defect with low recurrence risk⁴ (e.g. congenital diaphragmatic hernia).

IIIA - Disease with a known genetic predisposition, including with heritability (h^2) $\geq 50\%$ [e.g. asthma, autism spectrum disorder] *or* with Lambda s (λ_s) ≥ 10 [e.g. type I diabetes] ⁵

IIIB - Disease with multiple known causes, sometimes genetic, including those with $h^2 < 50\%$ *or* $\lambda_s < 10$ [e.g. atopic dermatitis]⁵

IV - Acquired disease with no or low genetic contribution (infection, trauma).

V - No disease.

¹ as per Gardner and Sutherland, 1996

² as per McKusick's Online Mendelian Inheritance in Man (OMIM™)

³ as per Friedman and Hanson 2002

⁴ as per Harper 2003

⁵ as per King et al. 2002

The final category for each separation was the lowest numerical assignment that appeared among the diagnoses list, and that was related to the patient's need to be hospitalized. Initially, a single research nurse classified the separations, however before a final category was assigned, one medical geneticist (thesis author BF) reviewed each diagnosis list. The geneticist used knowledge of the phenotype to make a subjective determination about whether or not the genetic diagnosis made an important contribution to the patient's need to be hospitalized. For example, if a patient with cystic fibrosis is admitted with pneumonia, a clearly causal relationship exists, and the separation was classified as IB (single-gene disorder). However, if a boy is admitted with gastroenteritis and happens to have a limb reduction defect which did not contribute to his need to be hospitalized, the final category was assigned as IV (acquired disease).

While most of the analysis was done using the "*final category*", in order to allow our data to be compared with some of the earlier studies, each separation was also assigned a "*hierarchical final category*" defined as the lowest numerical assignment that appeared among the diagnoses list, regardless of whether it contributed to the patient's need to be hospitalized.

3C. ANALYSIS: Most of the data analysis was performed using the "*final category*" of each hospitalization, as described above. Because of relatively small numbers in each of the final categories, these were stratified into three groups based on heritability. These groups were used for the hospital utilization analyses and are as follows:

- **Group 1:** separations with a **strongly genetic** final category, i.e. IA, IB, ID, IE.
- **Group 2:** separations with a **moderately genetic** final category i.e. IIA, IIIA.

- **Group 3:** separations with a **minimally or non-genetic** final category i.e. IC, IIB, IIIB, IV, V.

In order to allow our data to be compared to earlier studies of the burden of genetic disease among hospitalized children, it was reanalyzed in the following way. Each separation was assigned a “*hierarchical final category*” defined as the most strongly genetic diagnosis from the diagnoses list *regardless* of whether this diagnosis influenced the patient’s need to be admitted or the course of the hospitalization.

3C.1 Analysis of the validation study data: Over an 8-month period, 201 pediatric inpatients from the Janeway Hospital were randomly selected, and informed consent was obtained from the child’s parent or guardian. The admission was assigned a “final category” by the gold standard method (interview and record review by a clinical geneticist) and by the research nurse-method used in the broader study. The accuracy of the research-nurse method for assigning the final category (one of 11 choices) was calculated. The accuracy of this method for assigning separations to one of three genetic content groups (strongly, moderately or weakly genetic) was also calculated.

3C.2 Determination of the genetic content of 4,144 hospitalizations: The study’s Access[™] database was converted into an SPSS file. Analyses were performed on the SPSS file to determine the following:

- The total number of diagnoses for 4,144 consecutive hospitalizations, the average number of diagnoses per hospitalization and the proportion of the total diagnoses that fell into the strongly, moderately and minimally genetic groups (groups 1-3 above).

- For 4,144 consecutive separations, the frequencies in each of the 11 “final category” groups (IA, IB, IC etc), and the frequencies when these final categories were stratified into three groups based on the burden of genetic disease (groups 1-3).
- The analysis described in the previous bullet was repeated for 3,281 unique separations.
- To allow our data to be compared to earlier studies, each of the 4,144 separations was reassigned a “hierarchical final category” and the frequencies in the 11 genetic content groups (IA, IB, IC etc) were recalculated. The data analysis in the earlier pediatric hospitalization studies was mainly done using the total number of separations rather than just the unique ones.
- In order to look for trends supportive of the hypothesis that certain genetic problems are over-represented in the NL population, the 3,281 unique separations were analyzed in the two ways. Firstly, the frequencies of the most common specific diagnoses within four of the 11 “hierarchical final categories” were calculated: IA (chromosomal); IB (Mendelian disorders); IIA (high heritability birth defects); and IIIA (high heritability complex diseases). Secondly, for unique separations that were assigned to a IB “hierarchical final category”, the frequencies of the specific diagnoses grouped by modes of inheritance were determined.

3C.3 Analysis of hospital utilization data: As described in *section 3B.3*, the “final category” of each of the 4,144 consecutive hospitalizations was used to stratify them into three groups (strongly, moderately and minimally genetic). The total number of separations, rather than the unique number, was used because we were interested in

quantitating the net burden that children with genetic disorders place on the hospital system.

For each group, we calculated the mean length of stay (LOS) in days, along with the standard deviation (SD) and the 95% confidence interval (CI). The means between the three groups were compared using an ANOVA test.

We determined the sum of the person years in each of the three genetic content groups, and used this to calculate the cumulative surgeries per person years and the cumulative hospital days per person years for each of the groups. The rates between the three groups were compared using the chi-square test (6 x 6 cross table relationship).

3C.4 Multivariate regression to test the hypothesis that the genetic content of an admission influences the length of stay: Analysis of Covariance (ANCOVA) was used to test the main and interaction effects of three factors (including the “final category” of the separation which reflects its genetic content) plus one covariate (age at hospitalization) on length of stay (LOS). This model was selected because it allows use of both categorical and continuous variables and can incorporate categorical variables with multiple items (e.g. the final category was one of 11 items). The analysis was performed on 3,281 unique separations to avoid bias from repeat sampling.

The model’s dependent variable was LOS which is a continuous variable. The independent variables in the model were one covariate (age at admission, also a continuous variable) and three factors which are categorical variables: sex; “final category” of the separation; and the 20 most common “Case Mix Groups” (CMG’s).

The “Case Mix Group” code is further discussed below, but suffice it to say that the CMG factor reflects the main diagnosis that led to the patient’s requirement for hospitalization. It consists of 20 variables (the 20 most frequently occurring CMG’s in the dataset) which were compared to a reference variable (all other CMG’s not falling into the group of the 20 most frequent). The 20 most common CMG’s accounted for 1,759 of the 3,281 unique hospitalizations (53.6%).

Regarding the CMG codes, provincial hospital utilization data are collected and collated by the *Newfoundland and Labrador Centre for Health Information Services (NLCHI)*. These data are forwarded to the *Canadian Institute for Health Information (CIHI)*; each hospitalization’s CMG is assigned at the national level by CIHI. The “Case Mix Group” is the acute inpatient grouping that forms the foundation of much of CIHI’s statistical analyses. The CMG is determined by the patient’s “Most Responsible Diagnosis” (MRDx), coupled with any operative procedures that occurred during the hospitalization. In turn the MRDx is considered to be “*the one diagnosis that describes the most significant condition of a patient that causes his/her stay in hospital*”. The MRDx is used to assign the patient to one of 25 “Major Clinical Categories” (MCC) which identify either a body system (e.g. respiratory system) or other specific types of clinical problems (e.g. mental disorders, burns). Within each MCC, based on the presence or absence of an operative procedure, the case is directed towards a surgical or medical hierarchy. The CMG’s were originally numbered from 1-999, but subsequently many have been dropped, so that only 397 are currently in use. (CIHI website; personal communication with Ms. Kerry LeFresne, NLCHI, February 2006).

3C.5 Analysis of referral pattern data: This analysis was performed using 3,281 unique hospitalizations classified using the “hierarchical final category”, as described in section 3B.3. Each child was cross-referenced to the Provincial Medical Genetics Program’s database to determine if genetic services had ever been received. The proportion of patients in each category that had been referred was calculated. *A priori*, we decided that an acceptable clinical genetics referral rate for categories IA, IB, IC, ID, and IE was $\geq 80\%$.

We selected 60% as an acceptable referral rate for the IIA group (heritable birth defects). While selection of the cut-off was somewhat arbitrary, we decided that the IIA referral rate should be lower than the 80% cut-off for groups IA, IB, ID, IE where a genetics consult is essentially always indicated. Children with many of the IIA birth defects that are relatively prevalent in the general population (including spina bifida, cleft lip and palate) should be referred for genetic services. However, there are other IIA diagnoses which, while associated with an appreciable recurrence risk, do not otherwise warrant a genetics consult. This recurrence risk should ideally be reviewed with the parents although not necessarily by a genetics health professional. Examples of such IIA diagnoses include: some noncomplex forms of congenital heart disease including aortic stenosis and many septal defects; pyloric stenosis; hypospadias; and cryptorchidism.

CHAPTER 4:

RESULTS

4A RESULTS OF VALIDATION STUDY: As compared with the gold standard (geneticist interview), the discharge summary based method of assigning a final category had an overall accuracy of 94% (95% CI 0.90-0.97). 189 of 201 separations were assigned to the same final category (one of 11 categories IA-IV, as described in 3B.2) by both methods. **Twelve cases** were assigned to different final categories.

All 29 separations with a group 1 (strongly genetic) final category were correctly assigned to group 1 by the discharge summary method, which suggests that the larger study should identify almost all admitted children with a chromosomal or Mendelian disease that a clinical geneticist would recognize. Part of the reason that the detection rate for a group 1 diagnoses was so high is that category IE captures “syndromic” children with multiple congenital anomalies and/or developmental delay, even though a specific syndrome has not been diagnosed. Moreover the IE label can be applied based on the contents of the discharge summary even if the existence of a syndrome was not queried by the person that dictated the discharge summary. For 1/29 group 1 separations, the research nurse incorrectly assigned a 1E diagnosis (multiple birth defects suggestive of a syndrome) rather than a IB diagnosis (Mendelian syndrome) assigned by the clinical geneticist.

Of the 12 hospitalizations which were assigned to different final categories by the two methods, in only **6 cases** did this lead to a different group assignment (group 1, 2 or 3, see section 3C) as shown in the table below. For example, 61 cases were assigned to group 2 by the gold standard, and of these three were mislabeled as group 3 by the discharge summary method. Similarly of 111 hospitalizations assigned to group 3 by the

gold standard, three cases were incorrectly assigned to group 2 by the discharge summary method. Therefore the overall accuracy of the discharge summary method compared to the geneticist interview method for group assignment was $(201-6)/201 = 97\%$ (95% CI 0.94-0.99).

Table 4-1: Comparison of the categorization of 201 pediatric hospitalizations by discharge summary method compared with gold standard (review by a medical geneticist).

		Gold Standard -Interview			
Discharge summary		Gp1	Gp2	Gp3	Total
	Gp1	29	0	0	29
	Gp2	0	58	3	61
	Gp3	0	3	108	111
	Total	29	61	111	201

For the remaining **six discrepant cases**, the difference in “final category” did not result in a difference in the overall group classification, as shown below:

Table 4-2: *Number of separations that were categorized differently by the discharge summary method compared with the gold standard, but for which the miscategorization did not change the genetic content group.*

	Number of discrepant separations <u>within</u> group	Discrepancy (gold standard vs. discharge summary classification)
Group 1	1	IB vs. IE
Group 2	0	-
Group 3	5	IIB vs. IIIB IIIB vs. IV IV vs. IIIB IC vs. IV IIIB vs. IV
TOTAL	6	

From the validation study, we concluded that the discharge summary based method of determining the genetic content of an admission (as described in *section 3B.3*) was surprisingly accurate and therefore applied this method to the main dataset.

4B DETERMINATION OF GENETIC CONTENT OF 4,144 SEPARATIONS:

4B.1 Description of the distribution of the genetic content of 4,144 consecutive separations and 3,281 unique separations: The study population was defined as all pediatric patients admitted to the Janeway Hospital over a 14-month period (07/00-09/01). 4,144 consecutive pediatric hospitalizations were reviewed and categorized using the methods described in *section 3B.3*.

From the 4,144 separations, we retrieved 9,812 individual diagnoses, giving an average of 2.4 diagnoses per hospitalization. The frequency in each of the 11 categories of diagnosis is listed in *table 4-3*. When the individual diagnoses were categorized according to the burden of genetic disease, 3.8% were strongly genetic and 22.5% were moderately genetic (*table 4-4*).

Table 4-3: Category breakdown of 9,813 diagnoses for 4,144 consecutive patients.

Category of Diagnosis	Number of Diagnoses	Percent
IA	45	0.46
IB	183	1.86
IC	28	0.29
ID	12	0.12
IE	128	1.30
IIA	531	5.41
IIB	480	4.89
IIIA	1,681	17.13
IIIB	2,666	27.16
IV	4,004	40.80
V	54	0.55
TOTAL	9,813	100

Table 4-4: 9,812 diagnoses for 4,144 consecutive patients stratified into three groups based on the burden of genetic disease.

Genetic Burden	Category of Diagnosis	Number of Diagnoses	Percent
Group 1: Strongly Genetic	IA, IB, ID, IE	368	3.8
Group 2: Moderately Genetic	IIA, IIIA	2,212	22.5
Group 3: Minimally Genetic	IC, IIB, IIIB, IV,V	7,232	73.7
TOTAL		9,812	100

Then each of the 4,144 consecutive separations was assigned a “final category” defined as: the lowest numerical assignment (i.e. the most strongly genetic diagnosis) that appeared among the patient’s diagnoses list and that contributed to the patient’s need to be hospitalized or to the length or course of his/her stay (*table 4-5*).

33.2% of the 4,144 hospitalizations were attributed to either a strongly or moderately genetic disease (8.3% and 24.9% respectively). See *table 4-6*.

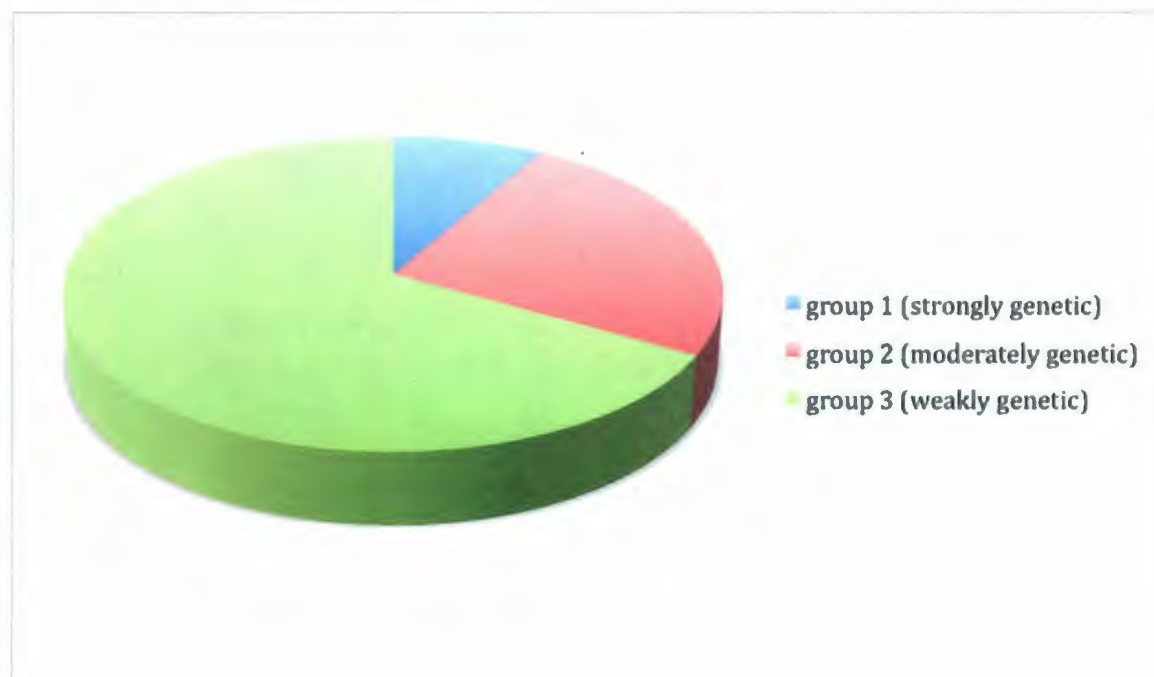
Table 4-5: Distribution of the “Final Category” for 4,144 consecutive hospitalizations.

Final category of hospitalization	Number of Hospitalizations	Percent
IA	39	0.94
IB	160	3.87
IC	22	0.53
ID	10	0.24
IE	133	3.21
IIA	220	5.32
IIB	150	3.60
IIIA	813	19.63
IIIB	1184	28.62
IV	1370	33.10
V	43	1.04
TOTAL	4,144	100

Table 4-6: “Final Category” for 4,144 separations, stratified into three groups based on the burden of genetic disease.

Genetic Burden	Final category of hospitalization	Number of hospitalizations	Percent
Group 1: Strongly Genetic	IA, IB, ID or IE	342	8.3
Group 2: Moderately Genetic	IIA, IIIA	1,033	24.9
Group 3: Minimally Genetic	IC, IIB, IIIB, IV or V	2,769	66.8
TOTAL		4,144	100

Figure 4-1: Proportion of 4,144 consecutive separations classified by 'Final Category' into three broader genetic content groups.



In order to compare our data to earlier studies of the burden of genetic disease among hospitalized children (see *Discussion, section 5B.1*), it was reanalyzed and each separation was assigned a “**Hierarchical Final Category**” defined as the most strongly genetic diagnosis from the separation’s diagnoses list *regardless* of whether this diagnosis influenced the patient’s hospitalization. See *tables 4-7, 4-8, 4-9 and 4-10*.

Table 4-7: *Distribution of “Hierarchical Final Category” for 4,144 consecutive admissions.*

“Most Genetic Final Category” of the hospitalization	Total number of Hospitalizations	Percent
IA	43	1.0
IB	173	4.2
IC	8	0.2
ID	12	0.3
IE	122	2.9
IIA	296	7.1
IIB	169	4.1
IIIA	971	23.4
IIIB	1,148	27.7
IV	1,160	28.0
V	42	1.0
TOTAL	4,144	~100.0

Table 4-8: *Distribution of “Hierarchical Final Category” for 3,281 unique admissions.*

“Most Genetic Final Category” of the hospitalization	Number of unique hospitalizations	Percent
IA	30	0.91
IB	120	3.66
IC	7	0.21
ID	10	0.30
IE	71	2.16
IIA	217	6.61
IIB	137	4.18
IIIA	778	23.71
IIIB	809	24.65
IV	1063	32.40
V	39	1.19
TOTAL	3,281	~100.0

Table 4-9: “Hierarchical Final Category” for 4,144 consecutive separations, stratified into three groups based on the burden of genetic disease.

Genetic Burden	Hierarchical final category of hospitalization	Number of hospitalizations	Percent
Group 1: Strongly Genetic	IA, IB, ID or IE	350	8.4
Group 2: Moderately Genetic	IIA, IIIA	1,267	30.6
Group 3: Minimally Genetic	IC, IIB, IIIB, IV or V	2,527	61.0
TOTAL		4,144	100

Table 4-10: “Hierarchical Final Category” for 3,281 unique separations, stratified into three groups based on the burden of genetic disease.

Genetic Burden	Hierarchical final category of hospitalization	Number of hospitalizations	Percent
Group 1: Strongly Genetic	IA, IB, ID or IE	231	7.0
Group 2: Moderately Genetic	IIA, IIIA	995	30.3
Group 3: Minimally Genetic	IC, IIB, IIIB, IV or V	2055	62.6
TOTAL		3,281	~100

4B.2 Description of the most frequent specific diagnoses associated with 4 of the 11

hierarchical final categories: The individual diagnoses within four of the 11 “hierarchical final categories” (IA, IB, IIA and IIIA) were reviewed, looking for trends that support the hypothesis that NL is enriched for particular genetic conditions. *Table 4-11* lists the most common specific diagnoses within each of the above four categories, as well as the number of unique cases associated with each diagnosis. The diagnosis frequencies were calculated as the number of unique individuals with that particular diagnosis divided by the number of unique patients within that category.

For example, during the 14-month study period, there were 43 hospitalizations that were assigned a hierarchical final category of IA (chromosomal disorder). However, 13 of these were repeat admissions.

Table 4-11: Description of the most frequent specific diagnoses associated with the "hierarchical final categories" IA, IB, IIA, and IIIA, for 3,281 unique separations.

Most Genetic Final Category	# of separations	# of unique separations	Specific diagnosis	# of unique cases	% within cat.
IA (chromosomal)	43	30	Down syndrome	14	46.7%
			Microdeletion syndromes	10	33.3%
			Microdeletion 22q11	4	13.3%
			Angelman / Prader Willi syndrome	4	13.3%
			Turner syndrome	2	6.7%
IB (single-gene)	173	120	Cystic fibrosis (AR)	17	13.4%
			Familial adenomatous polyposis (AD)	4	3.1%
			Factor V Leiden (AD)	4	3.1%
			I-Cell disease (AR)	3	2.4%
			Neurofibromatosis -1 (AD)	3	2.4%
			Noonan syndrome (AD)	3	2.4%
			Congenital adrenal hyperplasia (AR)	3	2.4%
			Metachromatic leukodystrophy (AR)	2	2.4%
IIA (heritable birth defects)	296	217	Congenital heart disease	52	24.0%
			Neural tube defect	39	18.0%
			Pyloric stenosis	27	12.4%
			Cleft lip &/or palate	25	11.5%
			Cryptorchidism/hypospadias	16	7.4%
			Microcephaly	13	6.0%
			Hearing loss (congenital with no environmental cause)	10	4.6%
			Club foot	9	4.1%
			Brain malformation	6	2.8%

Most Genetic Final Category	# of separations	# of unique separations	Specific diagnosis	# of unique cases	% within cat.
IIIA (heritable complex diseases)	971	778	Imperforate anus	6	2.8%
			Hirschsprung disease	4	1.8%
			Asthma	302	38.8%
			Seizures (recurrent & non-febrile)	80	10.3%
			Insulin dependent diabetes	76	9.8%
			Developmental delay	68	8.7%
			Depression	42	5.4%
			Attention deficit hyperactivity disorder	41	5.3%
			Crohn's disease + Ulcerative Colitis	36	4.6%
			Vesicoureteric reflux	35	4.5%
			Febrile seizures	31	4.0%
			Autism	15	1.9%
			Scoliosis	12	1.5%
			Juvenile rheumatoid arthritis	9	1.2%

For the specific diagnoses within the IB hierarchical final category, *Table 4-12* gives a breakdown of the modes of inheritance of these diagnoses, i.e. autosomal dominant; autosomal recessive; X-linked recessive; X-linked dominant; or mitochondrial. For some diagnoses, classification was not possible because the disease is genetically heterogeneous, with more than one possible mode of inheritance.

Table 4-12: *Distribution of the modes of inheritance for the specific single-gene diagnoses within the IB “hierarchical final category”, representing 127 unique hospitalizations.*

Mode of inheritance of IB specific diagnosis	Number of unique cases	Percentage of total single gene disorders
Autosomal dominant	51	40.2%
Autosomal recessive	50	39.4%
X-linked recessive	5	3.9%
X-linked dominant	1	0.79%
Mitochondrial	2	1.6%
Mixed modes of inheritance	18	14.2%
TOTAL	127	100%

4C THE IMPACT OF AN ADMISSION'S GENETIC CONTENT ON THE

UTILIZATION OF HOSPITAL RESOURCES: The "final category" of each of the 4,144 consecutive admissions was used to stratify them into three broader genetic content groups (strongly, moderately and weakly genetic). For each group, we calculated the mean length of stay (LOS) in days, the standard deviation (SD) and the 95% confidence interval (CI), as show in *table 4-13*.

The ANOVA test showed that there was a statistically significant difference in the mean LOS between at least two of the groups. When a Bonferroni correction was applied, the difference between all three pairs of groups was statistically significant.

Table 4-13: *Impact of the "final category" on length of stay (LOS) for 4,144 consecutive separations.*

	Strongly genetic = Group 1 (Final categories IA,IB,ID,IE)	Moderately genetic = group 2 (Final categories IIA, IIIA)	Minimally genetic = Group 3 (Final categories ID, IIB, IIIB, IV, V)	TOTAL
N (# of separations)	342	1,033	2,769	4,144
Mean LOS in days (SD)	8.01 (26.96)	6.37 (9.69)	3.99 (6.57)	4.92 (10.68)
95% CI	5.14 – 10.08	5.78-6.97	3.75-4.24	4.59-5.24

ANOVA: Group 1 versus group 2 $p = 0.04$

Group 2 versus group 3 $p < 0.001$

Group 1 versus group 3 $p < 0.001$

We determined the sum of the person years for each of the three groups as way of correcting cumulative surgeries and cumulative hospital days for the patient's age. We used the sum of the person years for each group to calculate the cumulative surgeries per person years and the cumulative hospital days per person years for each group. This is shown in *table 4-14*. The rates between the three groups were compared using the chi-square test (6 x 6 cross table relationship). For both cumulative surgeries and for cumulative hospital days, there was a statistically significant difference across the three genetic content groups.

Table 4-14: *Impact of "final category" on cumulative surgeries and cumulative hospital days, for 4,144 consecutive separations.*

	Strongly Genetic = Group 1 (Final Categories IA, IB, ID, IE)	Moderately Genetic = Group 2 (Final Categories IIA, IIIA)	Minimally Genetic = Group 3 (Final Categories ID, IIB, IIIB, IV, V)	TOTAL
N (# of separations)	342	1,033	2,769	4,144
Sum of person years	2,337	7612.89	20,154.52	
Cumulative surgeries / person years	706 / 2337 = 0.302	1,266 / 7,612.89 = 0.166	3,472 / 20,154.52 = 0.172	
Cumulative hospital days/ person years	31,809 / 2,337 = 13.61	27,552 / 7612.89 = 3.62	71,615 / 20,154.52 = 3.55	

Cumulative surgeries per person years: $\chi^2 = 160.546$

(p-value < 0.0001; correlation = .06705)

Cumulative hospital days per person years: $\chi^2 = 4002.292$

(p-value < 0.0001; correlation = 0.15571)

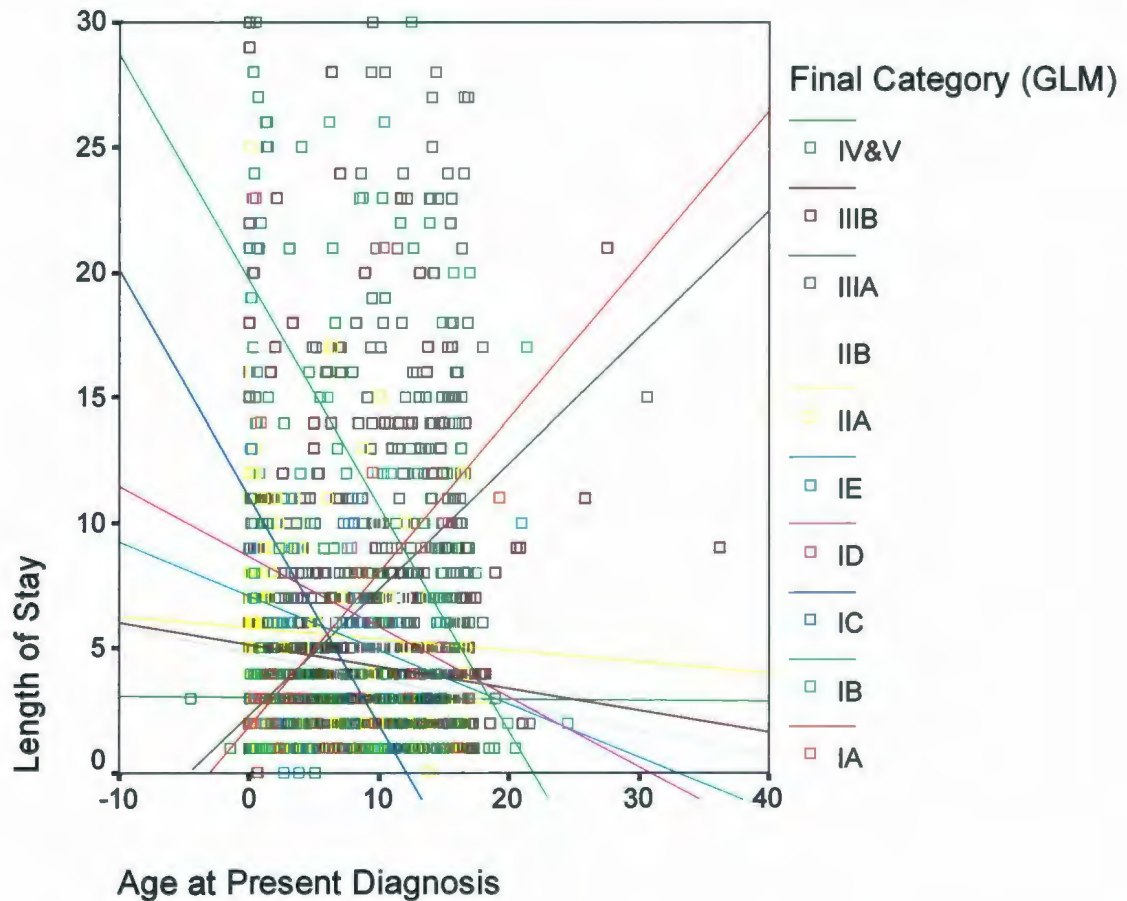
In order to allow our LOS data to be compared to the McCandless et al. 2004 study, we determined the mean age at admission and the mean LOS for our consecutive admission dataset by “hierarchical final category” (table 4-15). Children with Mendelian diseases (IB) and teratogenic exposures (IC) had the longest mean hospitalizations (10.6 d and 15.3 d respectively).

Table 4-15: Mean age and length of stay for 4,144 consecutive admissions by “hierarchical final category”.

Hierarchical final category	Number of separations	Mean age at admission (years)	95% CI for age (years)	Mean LOS (days)	95% CI for LOS (days)
IA	44	6.58	4.93 - 8.24	5.59	3.52 - 7.66
IB	172	8.61	7.69 - 9.52	10.57	5.03 - 16.11
IC	8	1.70	-0.26 - 3.67	15.38	-0.551 - 31.31
ID	12	9.20	6.20 - 12.21	5.58	2.01 - 9.15
IE	122	5.64	4.75 - 6.52	5.48	3.61 - 7.35
IIA	296	5.28	4.69 - 5.87	6.24	5.12 - 7.36
IIB	169	5.77	4.90 - 6.64	4.24	3.46 - 5.02
IIIA	971	8.09	7.74 - 8.44	6.40	5.78 - 7.02
IIIB	1148	8.15	7.82 - 8.49	4.41	4.04 - 4.78
IV	1160	7.19	6.88 - 7.50	2.97	2.69 - 3.25
V	42	0.47	-0.25 - 1.19	3.93	2.54 - 5.32
TOTAL	4,144	7.41	7.24 - 7.58	4.92	4.59 - 5.25

**This page is
missing from the
original book**

Figure 4-2: Test of the “homogeneity of regression slopes assumption” showing that there is a significant interaction between two of the model’s variables: “age at admission” and “final genetic content category”.



For example for children with a final category of IA (chromosomal) or IIIA (heritable multifactorial disease), LOS increases with increasing age, presumably reflecting the natural history of these conditions. Children with a final category of IB (Mendelian) or IC (teratogen) show the opposite pattern, with longer lengths of stay at younger ages. This makes intuitive sense because syndromic looking children generally have a chromosome analysis as a “first line” test (this may be ordered by the pediatrician prior to genetic consultation) so that such children rarely require early prolonged hospitalization in order

to reach a diagnosis. On the other hand, IB and IC children are frequently hospitalized early in life for a diagnostic work-up which involves consultation by several specialists, a dysmorphology assessment and genetic testing that cannot easily be ordered by a pediatrician. Therefore an interaction term (age x final category) was included in the model.

When the model was run, the following factors had significant p values of < 0.001 : *final category alone*; *CMG code*; and *final category x age*. Partial eta squared refers to the percent of the total variance of LOS that is explained by the variable which was 2.2% for *final category alone*; 3.6% for *CMG code* and 1.4% for *age x final category*. The p values for *age alone* ($p=0.226$) was not significant. In other words, when adjusted for the other variables (gender, final category, CMG code, and the interaction between final category and age), age alone was not a significant predictor of LOS. The p-value for *gender alone* ($p=0.111$) was also not significant.

Overall, the model was a relatively weak predictor of LOS. The significant variables (final category, CMG code, and age x final category) only explained 8.6% of the variance in length of stay.

The strength of the model improved when we removed the 21st CMG group and restricted the analysis to the 1,759 children captured by the 20 most common CMG codes. Then all variables apart from gender had significant p values: *final category alone* ($p=0.005$; 1.4% of variance in LOS); *CMG code* ($p < 0.001$; 25.4% of variance); *age alone* ($p=0.49$; 0.2% of variance); interaction term of *age x final category* ($p < 0.001$; 3.1% of variance). $R^2 = 0.338$ which means that the model explains ~34% of the variance in LOS, with clearly

the largest part of this attributable to the CMG code rather than to the genetic content of the admission.

The strength of the model did not improve when the final category was simplified as "IA (chromosomal) or IB (Mendelian)" versus "all other final categories". The "IA/IB group" contained 42 cases and the "other final category" group contained 1,717. The only variable with a significant p value ($p < 0.001$) was the CMG code which accounted for 30% of the variance of the length of stay. Neither the simplified final category or the "final category x age" interaction term had significant p values (0.154 and 0.72 respectively).

4E CLINICAL GENETIC REFERRAL RATES AMONG HOSPITALIZED

CHILDREN: This analysis was performed using the 3,281 unique admission dataset classified according to “hierarchical final category” because the presence of a genetic disease in a child is used to determine the need for a genetic consultation regardless of whether the genetic diagnosis was related to the child’s need to be hospitalized or to his/her hospital course. *A priori*, we decided that an acceptable rate for referral was \geq 80% for categories IA, IB, IC, ID and IE, and \geq 60% for category IIA (see *section 3C.5*). Each child was cross-referenced to the PMGP database to determine if genetic services had ever been received. The proportion of patients in each category who had been referred to the provincial genetics service was determined (*table 4-16*).

Table 4-16: Clinical genetic referral rates for 3,281 unique separations analyzed according to “hierarchical final category”.

Hierarchical Final Category of hospitalization	Number of patients referred / total number of patients	Percent of patients referred for genetic service	Target referral rate
IA	20/30	66.7	≥ 80
IB	90/120	75	≥ 80
IC	1/7	14.3	≥ 80
ID	8 / 10	80	≥ 80
IE	51/71	71.8	≥ 80
IIA	69/217	31.8	≥ 60
IIB	11/137	8.0	
IIIA	60/778	7.7	Selected cases (autism, developmental delay)
IIIB	27/809	3.3	-
IV	28/1063	2.6	-
V	0 / 39	0	-
TOTAL	365/ 3281	9.0	

CHAPTER 5:

DISCUSSION

5A VALIDATION STUDY: The discharge summary method of categorizing the genetic content of an admission was unexpectedly accurate compared with the gold standard geneticist interview. The accuracy for assigning one of 11 “final categories” to a hospitalization was 94%, with 97% accuracy for group assignment (strongly, moderately or weakly genetic). The narrow 95% confidence intervals for final category and group assignment (0.90-.97 and 0.94-0.99 respectively) indicate that the validation study had adequate power.

All the discharge summaries were reviewed by a single experienced research nurse (KW) who received careful hands-on training by a medical geneticist prior to beginning the study. The training resulted in a discharge summary method that was highly accurate when compared to a gold standard method. All children with a final diagnosis belonging to group 1 were correctly assigned to group 1 by the research nurse. Within group 1, there was one child with multiple congenital anomalies who was labeled IE (undiagnosed syndrome) by the research nurse and IB (specific single-gene disorder) by the geneticist, however this did not change the child’s overall group 1 assignment.

The validation dataset contains a higher proportion of strongly genetic cases than the full study dataset (*table 5-1*). The sample size for the main study was calculated to ensure that the smallest category from previous studies was represented. This was the chromosomal group at ~1% in the two most recent larger studies (McCandles et al. 2004; Carnevale et al. 1985). The validation study over-represented group 1 (strongly genetic) cases by over 2-fold. While the children used in the validation study were randomly selected, there may have been a bias towards the parents of children with strongly genetic conditions

consenting to participate (~20% of the families that were approached declined participation). Random stochastic effects observed with small samples are probably also contributing.

TABLE 5-1: *Comparison of final category group distributions between validation and main study datasets.*

	Group 1 Percentage (number of cases)	Group 2 Percentage (number of cases)	Group 3 Percentage (number of cases)	TOTAL
Validation study (gold standard)	14% (29)	30.3% (61)	53.7% (111)	100% (201)
Main study (unique separations)	6.7% (220)	25.3% (829)	68.0% (2232)	100% (3,281)

5B THE GENETIC CONTENT OF 4,144 CONSECUTIVE PEDIATRIC

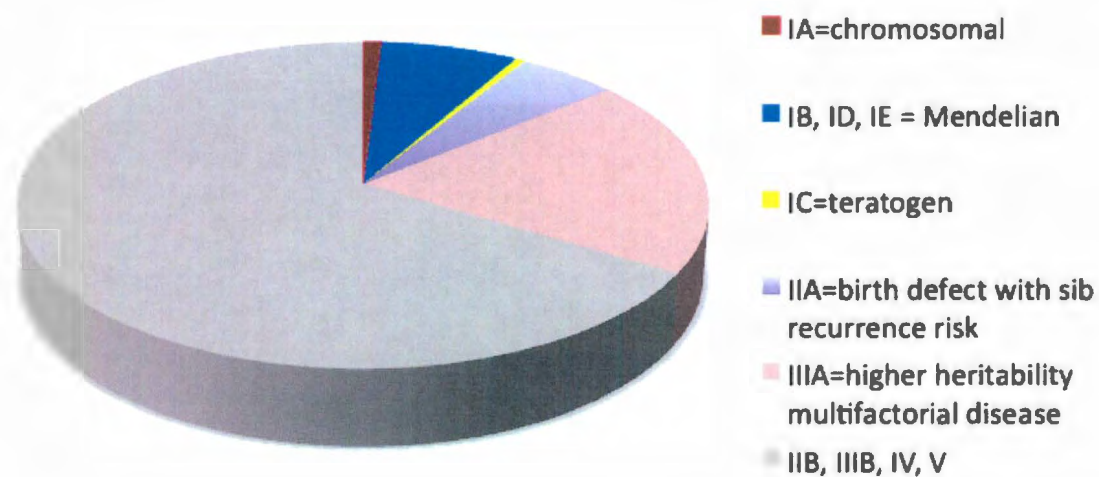
SEPARATIONS: *Tables 4-3 and 4-4* were included simply to show that most separations had more than one diagnosis (an average of 2.4 diagnoses per separation). From each diagnoses list, the one that had the highest genetic content and that was related to the patient's need to be hospitalized was selected, and this diagnosis was used to assign a "**final category**". Each diagnoses list was reanalyzed and the diagnosis with the highest genetic content was selected, regardless of whether it influenced the hospital stay. This diagnosis was used to assign each separation to a "**hierarchical final category**". The hierarchical classification allowed our findings to be compared to previous pediatric inpatient studies of the burden of genetic disease.

Tables 4-5 and 4-6 summarize the distribution of the final categories for the consecutive admissions dataset. One of the primary research questions was determination of the amount of genetic disease that exists in the pediatric inpatient population. This question was posed because we were interested in determining the burden that children with strongly genetic disorders impose on the health care system (hospitalization being one of the most expensive forms of healthcare). Hence it was appropriate to use total rather than unique hospitalizations and "final category" (i.e. the most genetic diagnosis that was related to the patient's need to be hospitalized).

The consecutive dataset contains 39 children with a chromosomal (IA) final category (0.94%) and 303 children with a definite (IB) or probable Mendelian (ID, IE) final category (7.4%). These two groups unequivocally require clinical genetic services (see *section 5E.1* for further discussion). Two hundred and twenty children (5.32%) had a

final category of IIA (heritable birth defect) for which genetic counseling or genetic testing is appropriate. Eight hundred and thirteen children (19.6%) had final category of IIIA (heritable complex disease). The only diagnoses within the IIIA group for which genetic services are presently indicated are autism and mental retardation/developmental delay (see *sections 5E.4.1 and 5E.4.2*). However over the next decade, genetic testing will probably be indicated for other IIIA diagnoses. This testing may be used to refine empiric recurrence risk counseling, to predict natural history and/or to guide therapeutic interventions. Complex disease genetic testing of this sort is likely to be largely ordered by non-geneticist physicians (for further discussion see *section 5E*).

Figure 5-1: Break down of 4,144 consecutive admissions by final genetic content category.



Approximately 1/3 of the children in the consecutive admission dataset had a strongly or moderately genetic disease, with the other 2/3 having a disease of low or minimal genetic

contribution (*table 4-6*). Our classification scheme was purposefully conservative. There is accumulating evidence that susceptibility to infections and to “sporadic” malignancies is genetically determined, but these diagnoses were classified at category IV (acquired) conditions.

5B.1 Comparison of the genetic content of the consecutive dataset to previous studies: The five earlier studies that examined the frequency and distribution of genetic diseases among hospitalized children all categorized the separations using the most genetic diagnosis regardless of whether it influenced the child’s hospitalization. Hence the consecutive admissions dataset was reanalyzed and categorized using the “hierarchical final diagnosis” (*table 4-7*).

The reanalysis resulted in small increases in the frequency of some of the strongly genetic diagnoses (e.g. IA went from 0.94% to 1%; IB increased from 3.9% to 4.2%), however this resulted in only a 0.1% increase in the percent of admissions that received a group 1 or strongly genetic categorization (to 8.4% from 8.3%). The moderately genetic group (group 2) increased from 24.9% to 30.6% at the expense of the minimally genetic group (group 3), which fell from 66.8% to 61% (*table 4-9*).

5B.1.1 Comparison of the genetic content of the consecutive dataset to previous studies - Limitations: Comparison of this study's frequencies to those of the earlier single hospital studies is limited by the following factors:

1. The five studies are spread over almost 40 years, over which time major advances have been made in our understanding of the genetic etiology of disease. Moreover, some formerly untreatable genetic diseases are now treatable and for other genetic diseases there has been substantial improvement in medical and/or surgical therapeutic options. These medical advances may have altered the amount and distribution of genetic diagnoses among hospitalized children.

The four earliest studies looked at pediatric inpatient populations from 1969-70 (Scriver et al. 1973), 1970 (Day and Holmes 1973), 1973 (Hall et al. 1978) and 1976-1980 (Carnevale et al. 1985). This leaves only the McCandless dataset (sample of inpatients from 1996) as one that was assembled reasonably close in time to our own (2000-2001).

By the 1970's, chromosome analysis had come into widespread use, however molecular genetic testing had almost no clinical utility (most of the ~2,000 genes associated with Mendelian diseases had not yet been identified). In spite of the testing limitations, the categorization of children into the broad disease categories used in all these studies (chromosomal, Mendelian syndrome, multifactorial birth defect etc.) would not have been substantially different in the 1970's compared with present day.

For example in the mid-1960's, a multifactorial inheritance model was put forward based on the observation that a large number of relatively common malformations (including

cleft lip and neural tube defect) clustered in families without conforming to the laws of Mendel. The model (Carter 1965) involved the concepts of genetic susceptibility (conferred by multiple genes) and a threshold for expression (determined by both genetic and environmental factors), and remains the basis of how we think about multifactorial diseases today. As a second example, the first edition of one of the most widely used texts books of dysmorphic syndromes, *Smith's Recognizable Patterns of Human Malformation*, was published in 1970 (Jones 2006).

However in comparing hospitalized children from the 1970's to today, one could reasonably expect shifts in disease distribution based on the introduction of screening programs and improved treatment options. For example newborn screening for PKU allows early diagnosis so that such a child may never require inpatient care. Newborn screening for tyrosinemia allows long-term survival of children with a formerly fatal condition. Surgical advances have been made so that hypoplastic left heart and congenital diaphragmatic hernia are no longer fatal in the newborn period, but these children require significant health care resources (including long term cardiology and surgical follow up).

2. The five earlier inpatient studies used much larger pediatric hospitals located in cities with populations of at least 1 million, compared to the Janeway Children's Health and Rehabilitation Centre which has 80 beds and is located in a city of 200,000. (table 5-3).

The Janeway Hospital handles 60% of all the pediatric admissions for a province of just over 500,000 people. Even if the Newfoundland pediatric population was enriched for genetic diseases requiring hospitalization, this might not be reflected in a comparatively higher frequency of strongly genetic admissions because these larger metropolitan

hospitals almost certainly admitted children from outside their catchment areas who required complex surgeries or multidisciplinary assessments not available through the local hospitals. The corollary is that some types of subspecialty surgeries are not available at the Janeway Hospital (including cardiovascular surgery, certain high complexity orthopedic and urogenital surgeries) so that children requiring these would have a proportion of their inpatient days spent at hospitals outside the province (primarily the Hospital for Sick Children in Toronto, the Shriner's institute in Montreal and the IWK hospital in Halifax).

3. No two of the studies used the same classification system for the genetic content of an admission, so that the only two categories that can be directly compared across all studies are the chromosomal and Mendelian groups.

For example Scriver et al. (1973) categorized all birth defects together rather than separating them into ones for which there is a sibling recurrence risk (e.g. congenital heart disease) and those which are of lower heritability and that have a low sibling recurrence risk (e.g. tracheoesophageal fistula). In addition, these authors did not distinguish between multifactorial diseases with low heritability (e.g. cancer) and more "purely" environmental conditions like infections and trauma. Like Scriver's group, Carnevale et al. (1985) classified all birth defects together. Also, the specific diagnoses for low heritability complex diseases (our category IIIB) and non-genetic diseases (our category IV) were not listed, so we cannot be confident that the group separated these two categories in the same way that we did.

Three of the earlier studies contained enough information about the specific diagnoses in each category to allow the published data to be reclassified using our genetic content categories: Day & Homes, 1973 (see *appendix C-1*); Hall et al. 1978 (see *appendix C-2*); and McCandless et al. 2004 (see *appendix C-3*).

The distribution across seven genetic content categories from the five earlier pediatric hospitalization studies and from this study is summarized in *table 5-2*. *Table 5-3* compares the proportion of group 1 and 2 separations (strongly and moderately genetic) across the six studies.

While Hall et al. used a classification system with some differences from ours, this was the only earlier study which listed the *majority* of the specific diagnoses within each genetic content category, along with the frequency of the diagnoses. This allowed Hall's data to be reanalyzed using our genetic content categories (see *appendix C-2* for details).

In terms of making direct comparisons with our study's frequencies, the McCandless paper (2004) was the most challenging. While the authors listed the specific diagnoses within each of their broader genetic categories (IA, IB, IC etc.) as an online appendix, the frequency with which each specific diagnosis occurred within the total number of hospitalizations was not reported, so that the McCandless dataset could not be reanalyzed using our classification scheme.

McCandless category IA included chromosomal and Mendelian conditions, 74 and 555 cases respectively (corresponding to our categories IA and IB). McCandless category IB (labeled "*multifactorial/polygenic diseases*") contained 832 cases or 14.5% of their

dataset and fairly closely corresponded to our category IIA of heritable birth defects. However, it also included the following diagnoses which we put into group IIIA (i.e. heritable non-congenital multifactorial diseases): depression, seizures, celiac disease and scoliosis. McCandless category IC (labeled "*heterogeneous causes, often genetic*") contained 495 separations or 8.7 % of their dataset. It contained a mixture of congenital and non-congenital disorders of both lower and higher heritability (i.e. a mixture of our categories IIA, IIB, IIIA and IIIB - see *appendix C-3* for the specific diagnoses).

McCandless category IIA ("*malformations of unknown etiology*") contained 48 cases or 0.84% of their total separations. It roughly corresponded to our low heritability birth defects (IIB), but as explained above, some of these were also included in the McCandless IC group. McCandless IIB (19 cases) completely corresponded to our category IC of teratogenic exposures.

McCandless category III ("*acquired disorders with genetic predisposition*") included 2,096 cases or 36 % of their total separations. This group was a mixture of our categories IIIA and IIIB (i.e. high and low heritability non-congenital diseases). McCandless category III included all pediatric malignancies which we categorized as IIIB (multifactorial diseases with lower heritability). McCandless categories IV and V overlapped with our categories IIIB, IV and V (i.e. all conditions with low or no genetic contribution).

See *appendix C-3* for a more detailed breakdown of the overlap between our genetic content categories and those of McCandless. Suffice it to say that the conclusion from the analysis (summarized in *table 5-2*) is that the only categories that can be compared

between the McCandless study and the present study are the chromosomal, Mendelian and non-genetic ones.

Table 5-2: *The distribution across seven genetic content categories for six pediatric hospitalization studies, including the present study.*

Study	Hospital	Sample size	Single gene (%)	Chromosomal (%)	High heritability birth defect (%)	Low heritability birth defect (%)	Complex disease with sib RR (%)	Complex disease with low RR (%)	Non-genetic (%)
Day and Holmes, 1973*	Massach. General Hospital	200	4.0	0	11	21.5 ^a	8.5	9	46.0 (includes malignancies apart from Wilms tumor)
Scriver et al. 1973	Montreal Children's Hospital	1,089	6.9	0.4	< 18.5 ^b	< 18.5 ^b	3.85 ^c	< 70.4 ^d (assumed to include malignancies)	< 70.4 ^d (assumed to include malignancies)
Hall et al. 1978*	Children's Orthopedic Hospital & Medical Center, Seattle	4,115	4.18	0.63	10.2	9.3	10.7	18.2 (includes malignancies)	46.6
Carnevale et al. 1985	National Institute Peds., Mexico	2,945 includes neonates	5.1 ^c	0.85	< 22.4 ⁱ	< 22.4 ⁱ	9.3	16.6 ^g	45.4 ^h
McCandless et al. 2004*	Rainbow Babies & Children's Hosp. Cleveland, OH	5,747	9.5	1.29	< 24 ⁱ	< 9.44 ^j	< 59.6 ^k	< 45.1 ^l	27.2
Present study	Janeway Hospital	4,144	7.4 (IB, ID, IE)	1.0 (IA)	7.1 (IIA)	4.1 (IIB)	23.4 (IIIA)	27.7 (IIIB)	29.0 (IV, V)

* see *appendix C* for details of how this study's dataset was reanalyzed using the genetic content categories from the present study.

^a 43 cases including 36 "urinary tract abnormalities", some of which we probably would have classified in our study as IIA (e.g. renal agenesis) or IIIA (e.g. vesicoureteric reflux), but the break down of individual diagnoses was not given.

^b Scriver et al. did not distinguish between high heritability congenital abnormalities (e.g. congenital heart disease) and low heritability birth defects (e.g. tracheoesophageal fistula).

^c In this group of 42 cases, there were 24 cases of "atopic hypersensitivity". The remaining 18 diagnoses were not specified.

^d Scriver et al. did not distinguish between low heritability complex traits and non-genetic diseases.

^eIncludes 101 children with an identified single-gene disorder + 50 children with multiple congenital anomalies (our study category IE).

^f Congenital defects were listed as broad categories (e.g. *congenital heart malformations*, *gastrointestinal anomalies*), so that this group could not be broken down into high and low heritability birth defects.

^g Includes 250 children with cancers/tumors + 37 with "*cerebral palsy and other neural system disorders*" + 37 with "*renal disorders*" + 156 with "*miscellaneous*" diagnoses.

^h This group consists of 1,341 children where the diagnoses were given as broad categories, e.g. "infectious diseases", "gastrointestinal", "nervous system". We likely would have categorized some of the specific diagnoses within each broad group as IIb and others as IV.

ⁱ Higher heritability birth defects appeared in three McCandless groups (IB, IC, IIA). The total number of separations in these three groups comprises 24% of their dataset, but these three groups also include non-congenital diagnoses and low heritability congenital anomalies.

^j Low heritability birth defects appeared in McCandless categories IC and IIA which together accounted for 9.44% of the dataset, but these two groups also included other diagnosis (that we categorized as IIA, IIIA, IIIB, ID, and IE).

^k Higher heritability non-congenital multifactorial diseases appeared in three McCandless groups: IB (14.5% of total separations); IC (8.6% of total separations) and III (36.5 % of total separations). However each of these three groups contained other diagnoses not falling within our IIIA group.

^l Lower heritability non-congenital multifactorial diseases appeared in McCandless categories: IC (8.6% of their total separations) and III (36.5% of their total separations). Both McCandless' groups contained other diagnoses not falling within our IIIB group.

Table 5-3: Comparison of frequencies across seven genetic content groups between Hall et al. 1978 and the present study – Chi-squared analysis.

Genetic content category	Hall et al. 1978 (n=4,115) Frequency (number of cases)	Present study (n=4,144) Frequency (number of cases)	Chi-squared
Single-gene	4.18 % (172)	7.4% (307)	$p < 0.0002$
Chromosomal	0.63% (26)	1.0% (43)	$p < 0.04$
High heritability birth defect	10.2% (421)	7.1% (296)	$p < 0.0002$
Low heritability birth defect	9.3% (381)	4.1% (169)	$p < 0.0002$
High heritability complex disease	10.7% (441)	23.4% (971)	$p < 0.0002$
Low heritability complex disease	18.2% (747)	27.7% (1148)	$p < 0.0002$
Non-genetic	46.6% (1916)	29.0% (1202)	$p < 0.0002$
Total	4,104 [excludes 11 teratogen/IC cases]	4,136 [excludes 8 teratogen/IC cases]	

Table 5-4: Comparison of six pediatric separation studies: type of hospital; sample size; method of categorization; and proportion of strongly genetic (group 1) and moderately genetic (group 2) hospitalizations.

Study	Type of hospital	Number of separations	Method of categorization	% with a group 1* final category	% with a group 2* final category
Day and Holmes 1973	Large metropolitan university pediatric and adult hospital (Boston)	200 admissions from 20 days over the year 1970 (? unique)	Review of "medical record".	4%	19.5%
Scriver et al. 1973	Large metropolitan pediatric hospital (Montreal) One of 2 pediatric hospitals for catchment area of 2.5 million	1,089 separations randomly sampled from 12,801 admissions (July 1969-July 1970) (not unique)	Extraction of one final diagnosis by chart review.	7.3%	< 22.4%
Hall et al. 1978	200-bed metropolitan pediatric university hospital (Seattle); Handles 55% of admissions for metro area; Total of 8,244 admissions in 1973	4,115 separations - all separations from alternate months in 1973 (not unique)	Review of discharge sheet (+ medical chart in >50% of cases). Assigned to a category by most genetic diagnosis even if it did not influence hospitalization.	5.1%	20.9%
Carnevale et al. 1985	350-bed pediatric university hospital (Mexico City); 7,000 admissions per year	2,945 separations - 1000 randomly admissions from 1976, 1978 and 1980 (not unique)	Review of discharge sheet (+ some charts). Category assigned using the most genetic diagnosis.	5.95%	< 31.7%
McCandless et al. 2004	244-bed pediatric university (Cleveland OH)	5,747 consecutive separations - every child admitted during 1996 (analysis performed on consecutive dataset)	Review of discharge code → discharge summary → full chart (abort only if Mendelian or chromosomal dx identified). Category assigned using most genetic diagnosis.	10.8%	< 83.6% (but includes cancers)
Present study	80-bed pediatric university hospital (St. John's NL)	4,144 consecutive; all separations from July 2000 - Sept 2001	Review of discharge codes + discharge summary (+previous discharge summaries)	8.4%	30.5%

* group 1 final category = chromosomal or single-gene disorder

* group 2 = heritable birth defect (e.g. cleft lip) or heritable complex disease (e.g. asthma)

5B.1.2 Comparison of the genetic content of the consecutive dataset to previous

studies - Discussion: As explained above, the three categories that can be best compared across all six consecutive admission studies are the chromosomal, Mendelian and non-genetic ones.

The frequency of **chromosomal disorders** ranged from 0-1.29%, and so was quite comparable across the studies. Only the McCandless study had a chromosomal disorder frequency that was higher than our own (1.29% versus 1%). Because most chromosomal disorders in live born infants are *de novo*, the incidence of these is not predicted to be higher in founder populations like that of Newfoundland.

One factor that could theoretically increase the prevalence chromosomal disorders in the NL population is the low uptake of prenatal diagnosis among pregnant Newfoundland couples. Only about 25% of pregnant women in Newfoundland opt for serum screening for Down syndrome (comparable to most rural Canadian populations) and we know that the majority of woman with a fetus with prenatally diagnosed Down syndrome terminate the pregnancy. Lower uptake of prenatal screening could in theory lead to a higher proportion of children that need hospitalization because of a chromosomal disorder. (Cavanagh et al. 2007).

Secondly, balanced chromosome rearrangements are found in 1 in 500 people. While these do not usually cause medical problems for the carrier of the translocation or inversion, that person is at risk of having chromosomally unbalanced offspring. Such a conceptus could either be miscarried or could be live born with psychomotor retardation and other medical problems caused by the unbalanced karyotype. Because of the large

sibships that were typical in Newfoundland families until this generation, a particular balanced rearrangement could become more common here than in a population with smaller family sizes. Dr. Penney Allderdice characterized one such rearrangement, a paracentric inversion of chromosome 3 which in the unbalanced form causes a severe psychomotor retardation syndrome called Allderdice syndrome. She recorded the outcomes of 54 pregnancies of inversion 3 carriers and 40% of these live born infants had unbalanced karyotypes. All the carriers of this inversion can be traced back to one particular NL couple (Allderdice et al. 1975; Nussbaum et al. 2001²). Because of the severe phenotype, virtually all carriers living in this province have opted for prenatal diagnosis (amniocentesis) and no baby with an unbalanced karyotype related to inversion 3 has been born in Newfoundland in over ten years (PMGP database).

Across the six studies, the frequency of **Mendelian diseases** ranged from 4% to 9.5%. Again only the McCandless study had a higher frequency of single-gene disorders than our own (9.5% versus 7.4%). Our original hypothesis was that because certain Mendelian diseases are overrepresented in the Newfoundland founder population, this might be reflected in a study of hospitalized Newfoundland children. As discussed above, the frequencies in the Mendelian category can be reliably compared across studies. The five earlier studies examined admissions to large metropolitan children's hospitals serving admixed populations and the Mendelian frequency in our study was not strikingly elevated in comparison with these.

There are a number of possible explanations. Perhaps there are over-represented Mendelian diseases among NL children, but these do not frequently lead to

hospitalization, for example syndromic and non-syndromic forms of mental retardation. Alternatively, the bulk of unusually common single-gene disorders in Newfoundland may be adult-onset conditions. Known examples include several autosomal dominant cardiomyopathies and hereditary cancer syndromes, including ARVD5, hereditary nonpolyposis colorectal cancer (HNPCC) syndrome, multiple endocrine neoplasia syndrome type 1 (MEN1) and hereditary diffuse gastric cancer syndrome, all of which only rarely manifest in children (Merner et al. 2008; Stuckless et al. 2007; Olufemi et al. 1998; Kaurah et al. 2007).

The McCandless study has the highest frequency of both chromosomal and Mendelian hospitalizations. This study's hospital was comparable to the hospitals of the other four earlier publications (metropolitan in location and serving an admixed population). The high chromosomal and Mendelian frequencies reported by this group may reflect availability of medical or surgical expertise at this particular hospital which led to admissions from outside the catchment area, i.e. enrichment for medically or surgically complicated children many of whom had an underlying genetic diagnosis.

Not only were strongly genetic diseases not over-represented in our sample of hospitalized NL children compared with children from more admixed populations, neither were isolated apparently multifactorial birth defects. None of the previous studies divided congenital anomalies into two groups based on relative heritability, so we can compare only the *total* birth defect frequencies. In our study, 11.2 % of the admitted children had either a high or a low heritability birth defect. Only three of the five earlier studies used a classification scheme that determined the frequency of children with non-

syndromic birth defects and these frequencies are as follows: 18.5% (Scriver et al. 1973); 19.5% (Hall et al. 1978) and 22.4% (Carnevale et al. 1985).

The one group of diseases that did appear to be more frequent in our dataset were non-congenital multifactorial diseases that either have a heritability of $\geq 50\%$ or a lambda $s \geq 10$. These conditions include asthma, autism, type I diabetes, inflammatory bowel disease and juvenile rheumatoid arthritis. High heritability (non-congenital) multifactorial diseases were present in almost one-quarter of our consecutive admissions (23.4%), compared with frequencies of 8.5-10.7% in three of the earlier studies (Day and Holmes 1973; Hall et al. 1978; Carnevale et al. 1985). Unfortunately the comparable frequency from the McCandless study, which is the most recent of the earlier studies, could not be determined. It is possible that the hospitals used in the three earlier studies had better resources for the outpatient treatment of chronic disease which reduced the need for hospitalization. Moreover the incidence of some of these IIIA diagnoses, for example autism and asthma, has increased worldwide over the past 30 years (Persico and Bourgeron 2006). However the comparatively high proportion of children in our study admitted with a IIIA diagnosis may also be due to increased disease prevalence attributable to Newfoundland's unique genetic architecture and/or to particular environmental factors acting on a genetically vulnerable background. We know for example that NL has one of highest incidences of type I diabetes in the world at 35 per 100,000 per year (Newhook et al. 2008).

The **non-genetic group** in our study can only be compared to two of the five earlier studies (Hall et al. 1978; McCandless et al. 2004), primarily because the other three

studies included children with malignancies in the non-genetic group. Hall's list of non-genetic diagnoses was very similar to our own (see *appendix C-2*) and her proportion of non-genetic admissions was 1.6-fold higher (46.6% versus 29%). One explanation is that over the almost 30 years between Hall's study and this one, there has been a shift towards outpatient management of less complicated pediatric problems (a higher proportion of which are non-genetic, for example milder infections, fractures). Alternatively, there may be a true difference in the proportion of pediatric hospitalizations attributed to wholly or partially genetic diseases between the populations. Unfortunately, these populations (Newfoundland versus Seattle) were studied 30 years apart and certain multifactorial pediatric disease have become more prevalent over this time frame. Nevertheless, when compared to Hall's study, the frequencies in all the genetic categories with the exception of birth defects were higher in our study— see *section 5B.1.3* for further discussion.

Interesting our non-genetic separation frequency (29%) was similar to that of McCandless (27.2%) suggesting that the drop in this category compared to Hall's study may be explained by a shift over time to handling of all but the most serious pediatric problems in an outpatient setting.

5B.1.3 Comparison of eight genetic content categories from this study's consecutive dataset to Hall et al. 1978: As detailed in *appendix C-2*, Hall's publication included frequencies of individual diagnoses which allowed her data to be reanalyzed using our classification scheme. Of the five earlier studies, this is the only one for which direct comparisons to our study can be made across all seven genetic content categories shown in *table 5-2*. When the two studies were compared, there was a statistically significant difference between six of the seven categories. Chi-squared showed a significant difference in frequencies for the following categories: single-gene; high heritability birth defects; low heritability birth defects; high heritability complex diseases; low heritability complex diseases; and non-genetic diseases (all p values <0.0002).

As mentioned above, any differences between the two studies may reflect changes that have occurred over the past 30 years in available treatments coupled with a trend toward more outpatient management or, alternatively these may be indicative of true differences in the burden of genetic disease between the populations (or both). Also certain higher heritability multifactorial diseases (e.g. autism, asthma) are either more commonly diagnosed or have genuinely become more prevalent than in the 1970's.

Compared with Hall's study:

- Our total birth defect proportion is *lower* (11.2% versus 19.5%).
- Our proportion of high heritability (non-congenital) multifactorial diseases is *higher* (23.4% versus 10.7%).
- Our proportion of lower heritability (non-congenital) multifactorial diseases which includes pediatric malignancies is *higher* (27.7% versus 18.2%).

- Our proportion of non-genetic admissions is *lower* (29% versus 46.6%).

5B.2 The most frequent Mendelian and chromosomal diagnoses for the 3,281 unique separations:

5B.2.1 The commonest Mendelian Diagnoses for 3,281 unique separations:

Approximately 6,000 Mendelian disorders have been described and about half of these are autosomal dominant (AD). Autosomal recessive (AR) disorders are the next most common accounting for about 1/3, with the most of the remainder (several hundred diseases) being X-linked recessive (XLR) (Nussbaum et al. 2001³).

In our dataset, there were 120 unique separations that were assigned a hierarchical final category of IB (single-gene diseases) and of these, there were about equal numbers of autosomal dominant and autosomal recessive disorders (40.2% and 39.4% respectively). Another 3.9% of the separations were X-linked recessive. In Hall's 1978 study, among consecutive admissions (not unique), there was a higher proportion of autosomal recessive and X-linked recessive disorders than in our study: AD 29.6% (commonest diagnosis: Neurofibromatosis type 1); AR 56.8% (commonest diagnoses: cystic fibrosis and sickle cell anemia); XLR 13.5% (commonest diagnosis: hemophilia). These findings are consistent with the hypothesis that a significant part of the burden of genetic disease in this province is due to autosomal dominant disorders which have become prevalent due to high sibship sizes, rather than due to AR diseases which collectively are still fairly rare.

The single commonest Mendelian disease among our 120 IB cases was also cystic fibrosis (17 unique cases accounting for 13.4% of the Mendelian phenotypes). CF is the commonest autosomal recessive disease in Caucasian children with an incidence of 1 in 2,000 in populations of Northern European ancestry. The carrier frequency in these populations is high (1 in 22) and the parents of most CF patients are not consanguineous (Nussbaum et al. 2001⁴). Because of its prevalence and the fact that affected children frequently need hospitalization, one would expect this to be a common diagnosis among hospitalized children drawn from any admixed North European population.

One of the next most common AR diagnoses was I-Cell disease (three cases). There have only been a few estimates of the incidence of this lysosomal storage disorder (1 in 252,500 live births in Japan; 1 in 625,500 in the Netherlands), but it is considered to be a very rare disorder in most populations (Leroy 2007; Poorthuis et al. 1999). In total 19 Newfoundland children have been diagnosed with I-cell disease (PMGP database) and all come from the same area on the Avalon Peninsula. Two patients have been genotyped and both have homozygous mutations in the *GNTFAB* gene so this is presumably a founder allele.

5B.2.2 The commonest chromosomal diagnosis for 3,281 unique separations: The vast majority of chromosomal abnormalities among liveborns are *de novo* events, so that one would not expect the prevalence of these to be increased in Newfoundland on the basis of this province's population structure. Among our unique dataset, 43 separations (0.91%) received a hierarchical final diagnosis of IA (chromosomal disorder) which is in keeping with the earlier pediatric hospitalization studies.

The overall incidence of chromosomal disorders in live born children is approximately 1 in 160 (0.7%) (Nussbaum et al. 2001⁵). This number does not include microdeletion /microduplication syndromes which were also included in our IA category. With the recent introduction of genomic microarray testing, we have learned that these “genomic” disorders are at least as common as the cytogenetic disorders identified by standard karyotyping (i.e. aneuploidy and structural rearrangements). This is further addressed below, but suffice it to say that microarray testing was not routinely ordered in 2000/2001 when the hospitalizations in this dataset occurred.

Of the 0.7% of liveborns with chromosomal abnormalities that can be identified with a standard karyotype, the commonest group of disorders are the sex chromosome aneuploidies which include Klinefelter syndrome (47,XXY) and Turner syndrome. Numerical abnormalities of the sex chromosomes (X and Y) are found in 1 in 360 male births and 1 in 580 female births. Within this group, only Turner syndrome can be diagnosed based on clinical features either in the neonatal period or before puberty. The next most common group of liveborns with chromosomal abnormalities that can be identified on standard karyotyping are the autosomal aneuploidies (1/700 live births) and the commonest of these is trisomy 21 (1 /830), which is the karyotype in 95% of cases of Down syndrome (Nussbaum et al. 2001⁵).

Within our unique dataset of 43 IA admissions, 14 (46.7%) had Down syndrome (DS) which was the most frequent diagnosis, followed by two cases of Turner syndrome (6.7%). In the McCandless study (2004), there were 74 admissions of patients with

chromosomal disorders (this was a consecutive admission data set); 69% had Down syndrome and 2.7% had sex chromosome abnormalities.

About half of children with DS have at least one major congenital malformation (the commonest being congenital heart disease) and which often leads the child to require hospitalization. Moreover DS is the commonest human malformation syndrome which almost never goes undiagnosed, so it is not surprising that this was the commonest chromosomal diagnosis in the dataset. Even though the sex chromosome aneuploidies are commoner than DS, these are generally diagnosed later in life (sometimes not until adulthood when the affected individual is investigated for infertility). Moreover they are associated with less need for hospitalization, although some individuals require gonadectomy.

Microdeletion syndromes accounted for 18/43 unique IA separations (41.9%). The two commonest of these were microdeletion 22q11.2 and Angelman/Prader Willi syndromes. These conditions (and all the microdeletion syndromes in the dataset) can usually be diagnosed based on clinical features, with follow-up confirmatory fluorescent *in situ* hybridization (FISH) testing which has been routinely available since the early 1990's. The dataset didn't include any microduplication syndromes or any of the more recently delineated microdeletion syndromes that are usually diagnosed by microarray testing.

5B.2.3 Commonest IIA diagnoses (heritable birth defect) and IIIA diagnoses (heritable non-congenital complex disease) for 3,281 unique separations: Our commonest IIA diagnosis was congenital heart disease (52 / 217 IIA admissions or 24%)

and our commonest IIIA diagnosis was asthma (302/778 IIIA admissions or 38.8%). These were also the most common diagnoses in the comparable groups from Hall et al. 1978: congenital heart disease accounted for 162/421 of her IIA admissions (38.4%) and allergy/asthma accounted for 157/441 of her IIIA admissions (35.6%). We were unable to correct the Hall dataset for unique hospitalizations. The authors reviewed 4,115 separations representing all hospitalizations from alternate months in 1973.

5C IMPACT OF THE GENETIC CONTENT OF AN ADMISSION ON THE UTILIZATION OF HOSPITAL RESOURCES: We hypothesized that children with strongly genetic disorders collectively utilize more hospital resources than children with non-genetic conditions. Our analysis of length of stay (LOS), cumulative surgeries and cumulative hospital days supported this hypothesis. We divided the 4,144 consecutive separations into three genetic content groups for all three analyses (group 1=strongly genetic; group 2=moderately genetic; group 3=minimally genetic).

There was a statistically significant difference in mean LOS between all three groups, with the group 1 children having the longest mean LOS (8.01 d) and the group 3 children having the shortest (3.99 d). The 95% confidence intervals of groups 1 and 3 did not overlap (*table 4-13*). Hence children with chromosomal or Mendelian disorders had an average hospital stay that was twice as long as children with minimally genetic conditions. This is consistent with the findings of McCandless et al. 2004 who reviewed 5,747 consecutive separations from 1996 to a pediatric hospital in Cleveland OH. Children with a single-gene or chromosomal diagnosis had an average hospital stay (7.1 days) that was twice as long as the average stay for children admitted with infections or

acute trauma (3.5 d). For the categories that could be directly compared, the mean ages at admission also seemed comparable between the McCandless study and our own. For example, the mean admission ages for our chromosomal and single-gene patients were 6.58 and 8.61 years respectively (*table 4-15*) and in the McCandless study, the average age across both groups was 7.7 years. Our average age at admission in the non-genetic category was 7.19 years (*table 4-15*) compared with 6.8 years in the McCandless study.

In our consecutive admission dataset, there were 342 group 1 (strongly genetic) separations (*table 4-13*). About half of these (160 hospitalizations) were children with single-gene disorders (IB), with the other half (182 separations) being children with chromosomal disorders (IA) and other syndromic children where the precise genetic mechanism is unknown (ID, IE). See *table 4-5*. Within group 1, the IB children had an average LOS (10.57 d) that was 50% longer than that of other group 1 disorders (5.48-5.59 d). However the longest average LOS (15.38 d) belonged to the IC separations, of which there were only eight (*table 4-15*). These are children with teratogenic exposures (e.g. alcohol, cocaine) where there are often social factors that necessitate hospitalization beyond the length of time needed for a diagnostic work up.

In order to compare cumulative surgeries and cumulative hospital days between the three groups, we first calculated the total number of person years for each group. Then, we divided the cumulative surgeries and hospital days within each group by the number of person years in that group (*table 4-14*).

Cumulative surgeries per person years for groups 1, 2 and 3 respectively were as follows: 0.302, 0.166, and 0.172. The chi-squared test gave a p-value of < 0.0001 which indicates that there is a significant difference between the rates of at least two groups. By inspection, the cumulative surgery rate between at least groups 1 (strongly genetic) and 2 (moderately genetic) must be significant. Put another way, there were 1.8 fold more surgeries in the strongly genetic group than in the moderately genetic group. The number of surgeries was slightly higher in the moderately genetic than the minimally genetic group. The moderately genetic group included children with heritable birth defects (IIA) that often require corrective surgery, but three quarters of the group were children with IIIA diagnoses (heritable non-congenital multifactorial diseases like asthma) most of which do not routinely require surgery (*tables 4-5 & 4-6*). The minimally genetic group included children with trauma which probably explains why group 3's cumulative surgery rate was slightly higher than the group 2 rate.

Cumulative hospital days per person years for groups 1, 2 and 3 respectively were: 13.61, 3.62 and 3.55. The chi-squared test was also statistically significant (p-value < 0.0001) so that by inspection the rates between at least the two most divergent groups (groups 1 and 3) must be significant. When corrected for age, group 1 children had 3.8-fold more hospital days than children in group 3. Hall et al. reviewed 4,115 consecutive admissions from alternate months in 1973 to a Seattle children's hospital and found the same trend. The authors looked at the average number of admissions in each of their genetic content groups and their "clearly genetic" group (chromosomal and Mendelian disorders) had an average of 5.3 admissions, more than 3-fold higher than patients with non-genetic disorders. However the average admission number was not corrected for age and the

mean age of the genetic patients (11.6 years) was significantly higher than that of the non-genetic group (8.6 years).

5D ANCOVA REGRESSION MODEL: We hypothesized that pediatric patients with genetic diseases utilize more hospital resources than children with non-genetic conditions and our analysis of mean LOS, cumulative surgeries and cumulative hospital days across three genetic content groups supported this hypothesis. Moreover our LOS data were very comparable to the most recent previous pediatric hospitalization study (McCandless et al. 2004). Hence we decided to use multivariate regression to examine the impact of the genetic content of an admission on length of stay (LOS). We hoped to be able to construct a model that would use several variables (including the genetic content of the admission) to identify patients at risk for prolonged length of stay.

ANCOVA was used to test the effects the following variables on LOS: final genetic content category; Case Mix Group code; sex; and age at admission. Because there was a significant interaction between two of the independent variables (age and final category) an “age x final category” interaction term was included in the model.

In the first regression analysis, all 11 genetic content final categories were used (IA-V) and the CMG codes were analyzed as the 20 most frequent CMG codes (which accounted for 53.6% of the 3,281 unique separations) with the remaining admissions grouped together as a 21st CMG code. This model explained only 8.6% of the variance in LOS with 3.6% attributable to the genetic content of the admission; final genetic category

alone explained 2.2% of the variance and “final genetic category x age” explained 1.4% of the variance in LOS.

In an attempt to improve the strength of the model, we removed the 21st “other” CMG category and analyzed only the 1,759 unique separations captured by the 20 most frequent CMG codes. This model explained 34% of the variance in LOS, but the majority of this (25%) was due to the CMG code, i.e. the diagnosis that led to the patient’s need to be hospitalized. The percent of the variance in LOS that was related to the admissions genetic content rose only minimally to 4.5% (final genetic category alone 1.4%; age x final category 3.1%).

We modified the analysis a third time by simplifying the final genetic category variable as “IA or IB” (42 cases) versus “all other genetic content final categories” (1,717 cases). Only the CMG code had a significant p-value and explained 30% of the variance in LOS.

Hence in spite of the fact that the mean LOS was significantly different over three genetic content groups, the bulk of the variance in length of stay in our unique admission dataset was determined by the admitting diagnosis, rather than by the presence or absence of a strongly genetic disease. For example, the above analysis suggests that the admitting diagnosis of pneumonia is a stronger determinant of LOS than the fact that the child has an underlying diagnosis of cystic fibrosis. Nevertheless from our data we can conclude that children with single-gene disorders (and those with teratogenic exposures) are at risk for prolonged lengths of stay (> 7days) (*table 4-15*).

The model did not incorporate a variable that reflects the place of residence of the child (for example St. John's area, the remainder of the Avalon Peninsula, outside the Avalon Peninsula). This factor also likely influenced LOS, with a tendency towards longer lengths of stay for children who live outside the St. John's area.

5E RATES OF REFERRAL FOR CLINICAL GENETICS SERVICES AMONG HOSPITALIZED CHILDREN:

A priori, we decided that 80% was an acceptable referral rate for children with the following diagnoses anywhere in their diagnosis list: IA (chromosomal disorders); IB (single-gene disorders); IC (teratogenic exposures); ID (genetic syndromes for which the precise genetic cause has not been identified); IE (syndromic children with no specific syndrome identified); and IIA (heritable birth defects). The Provincial Medical Genetics Program's database was reviewed in 2005, so that any genetics referral that was initiated during a 2000 or 2001 hospitalization would have been captured (either as a pending or completed consultation or as an appointment that was not kept).

5E.1 Referral rates for children with a strongly genetic diagnosis (IA, IB, ID, IE):

The 80% referral rate was only achieved for the **ID group** (known genetic syndrome, but precise genetic etiology not identified). Eight of 10 children were referred. The diagnoses of the referred were: three cases of Goldenhaar syndrome; two cases of VACTERL association and one case of VECS. One patient with Goldenhaar syndrome and one with Kabuki syndrome were not referred. These conditions are all characterized by a high degree of dysmorphism and/or multiple birth defects which probably explains the high referral rate.

Children who had a diagnosis that was either **IB** (chromosomal) or **IE** (multiple medical problems highly suggestive of a chromosomal or Mendelian disorder) also had reasonable referral rates of 75% and ~72% respectively.

When the individual diagnoses of the **IB** "referred" versus "not referred" children were reviewed, some trends were observed. For example, there were 24 children with *inborn errors of metabolism* and the majority (19 patients or ~79%) were referred for genetic services. The diagnoses of the referred included Leighs disease, infantile NCL, various amino and organic acidopathies and storage disorders, including three cases of I-cell disease. However the severity of these metabolic disorders as a group was not subjectively different from the diagnoses of the five metabolic cases that were not referred (one case of each of the following: Krabbes leukodystrophy; MERRF; metachromatic leukodystrophy; Pelazieus-Merzbacher disease and porphyria). Most of the children in the non-referred group would have presented with neurologic symptoms and may have been diagnosed by a pediatric neurologist.

Within the **IB** group, there was a trend towards under-referral of children with *Mendelian blood disorders*. Only two of nine were referred; their diagnoses were von Willebrand disease (VWD) and hemophilia. The diagnoses of those not referred included three cases of hereditary spherocytosis, one case of G6PDH deficiency and another case of VWD. Hereditary blood disorders are diagnosed by hematologists, with the role of the clinical geneticist/ genetic counselor being limited to genetic counseling +/- confirmation of diagnosis by arranging molecular genetic testing. This is in contrast to most other Mendelian conditions (for example dysmorphic syndromes, syndromes characterized by

multiple birth defects and hereditary cancer syndromes) where the primary diagnosis is generally made by a clinical geneticist.

There were six children with *Mendelian cancer syndromes*, of whom four were referred: two cases of familial adenomatous polyposis (FAP); one case of Li-Fraumeni syndrome; and one patient with multiple endocrine neoplasia type 1 (MEN1) syndrome. Two cases of FAP were not referred.

Fifteen of 17 children with *cystic fibrosis (CF)* were referred (~88%). Genetic testing is an important component of the diagnostic algorithm for a patient with symptoms compatible with CF and clinically relevant genotype-phenotype correlations exist. *Eastern Health's* molecular diagnostic service is organized so that (with the exception of testing for hereditary hemochromatosis and Factor V Leiden) all Mendelian genetic testing can only be ordered by one of the PMGP's clinical geneticists, which probably explains the high referral rate for CF patients.

Finally within the IB group, there are several examples of diagnoses for which some children with that specific diagnosis were referred, whereas others were not. These conditions included Noonan syndrome, congenital muscular dystrophy, Factor V Leiden, pseudocholinesterase deficiency, and VWD. In all these cases, a general pediatrician or pediatric specialist could reasonably have made the diagnosis. For example, Noonan syndrome could be diagnosed by a pediatric cardiologist, muscular dystrophy by a pediatric neurologist etc. The majority of the IB children who were not referred were likely diagnosed by a pediatrician although some may have been diagnosed by a geneticist outside of the province (either as part of an external consultation that was not

organized by one of the PMGP's geneticists, or because the family moved to NL and had a genetics consultation in another province prior to relocating).

Separations were assigned a IE diagnosis because the information contained within the discharge summary suggested to this study's research nurse and medical geneticist that the child was syndromic even though a specific syndrome was not identified. Therefore it is not surprising that these children had a fairly high referral rate (71%).

Twenty of 30 children with a IA diagnosis (~67%) were referred. There was no obvious difference between the diagnoses of the "referred" versus the "not referred" with respect to disease severity. The ten cases that were not referred included chromosome disorders that are identified by traditional karyotyping (five cases of Down syndrome, one child with a ring chromosome and one with trisomy 18) and also included microdeletion syndromes identified by fluorescent *in situ* hybridization (two cases of microdeletion 22q11.2 and one case of Angelman syndrome).

Down syndrome is the commonest chromosomal disorder among liveborns (as well as the commonest human malformation syndrome) and most pediatricians are very familiar with the phenotype. The *American Academy of Pediatrics* and others have published guidelines for anticipatory care of children with Down syndrome (*American Academy of Pediatrics* 2001; van Cleve and Cohen 2006). This probably explains why five of 13 children with Down syndrome were not referred for genetic services. Moreover, within the *Eastern Health* care system, any physician can order chromosome analysis and FISH testing, so genetic testing within the IA group is not contingent upon making a genetic

referral. This is in contrast to the IB group where genetic testing cannot occur without a genetics referral.

5E.2 Referral rates for children with a IC diagnosis (teratogenic exposure): There were seven children in the dataset with a IC diagnosis and only one (maternal HIV) was referred for genetic services. The non-referred included two children with fetal alcohol spectrum disorder, two infants of diabetic mothers and two children who were exposed to teratogenic medications. The referral pattern presumably reflects the fact that pediatricians are comfortable diagnosing these conditions.

In 2005 Chudley et al. published guidelines for the diagnosis of fetal alcohol spectrum disorders (FASD) in the *Canadian Medical Association Journal*. The article includes instructions about how to assess patients for dysmorphology typical of FASD and was published in a journal with readership that is much broader than the clinical genetics community. It emphasizes that FASD is a diagnosis of exclusion that cannot be confirmed by a specific biologic assay. FASD has a differential that includes a number of genetic syndromes that cannot be easily diagnosed by a non-geneticist so that referring these children for a genetic consultation is of value.

5E.3 Referral rates for children with a IIA diagnosis (high heritability birth defect):

The second largest referral deficit occurred for hospitalized patients with a IIA diagnosis as the most genetic condition with the diagnosis list. *A priori*, we suggested that a 60% referral rate might be appropriate (i.e. some rate < 80% referral rate cut-off for groups IA, IB, IC, ID and IE where a genetic consult is essentially always indicated). This is because

there are some IIA diagnoses which have an appreciable sibling recurrence risk, but do not otherwise warrant a genetic consult. This recurrence risk should ideally be reviewed with the parents although not necessarily by a genetic health professional. Examples of such IIA diagnoses include some noncomplex forms of congenital heart disease (e.g. aortic stenosis and septal defects), pyloric stenosis, hypospadias and cryptorchidism.

However even using this lower cut-off, children in the IIA group were under-referred. Only 69 of 217 children with high heritability birth defects (~1/3) were referred for genetic services.

Within the IIA group of 217 unique admissions, the referral rates for some of the commonest diagnoses were as follows:

- i. Congenital heart disease: 6/31 were referred (19%). Within this group there were six children with *complex* congenital heart disease of whom three were referred. Because cardiac surgery cannot be performed at the Janeway hospital, children requiring surgery are referred out of the province. Hence it is possible that some children in this group who needed cardiac surgery were seen by a geneticist in another province.
- ii. Spina bifida: 21/30 were referred (70%). Genetic consultation is indicated here to exclude rare syndromic causes of NTD, to review the recurrence risk for future pregnancies and prenatal diagnosis options, and to ensure that the mother is aware of the higher preconceptional folate supplementation recommendation (5 mg daily) for subsequent pregnancies (Wilson et al. 2007).
- iii. Cleft lip and/or palate: 18/22 were referred (82%). The high referral rate in part reflects the fact that these children receive outpatient medical care through the

Janeway's multidisciplinary craniofacial clinic and a genetic counselor from the PMGP is part of the craniofacial team.

The components of a genetic consultation for a child with a IIA diagnosis include:

1. A 3-generation pedigree, a pregnancy history to exclude teratogenic exposures, a medical history and examination of the child including an assessment for dysmorphic features. Depending on the findings, diagnostic imaging may be ordered to screen for other birth defects and to help exclude the existence of a broader genetic syndrome. As an example, while most clefts of the lip and palate are isolated birth defects, this malformation is an important feature of over 20 Mendelian and chromosomal disorders (Harper, 2003³). If such a child is not referred for genetic services, work up to exclude a more strongly genetic diagnosis may be arranged by the child's general or subspecialty pediatrician.
2. Genetic testing for selected cases. Examples include: FISH for microdeletion 22q11.2 for children with complex congenital heart disease or with cleft palates; chromosome analysis followed by possible single-gene testing for children with holoprosencephaly (HPE); molecular testing of the *RET* gene in a child with non-syndromic Hirschsprung disease; and testing for the Meunke syndrome recurrent mutation in the *FGFR3* gene in a child with apparently non-syndromic craniosynostosis. *The unique dataset included two children with HPE, one of whom was referred for genetic services and four cases of Hirschsprung disease none of whom were referred.*
3. Based in the findings in steps 1 and 2, recurrence risk counseling. Assuming that a broader syndrome is not identified, the parents of a child with a IIA isolated birth

defect have an empiric risk of having another affected child of 2-10% depending on the birth defect (Harper, 2003²).

The *negative repercussions* of not referring a child with a IIA diagnosis include:

1. *Potentially misclassifying a birth defect as isolated when it is a feature of a broader genetic syndrome.* If a child with an apparently isolated birth defect is carefully reviewed by a pediatrician this is unlikely to occur. While a pediatrician may not be able to provide a specific syndromic diagnosis, he or she is likely to be able to identify that the child's phenotype puts the patient into a syndromic category.

Nevertheless, non-referral could in some instances result in failure to recognize that the child has a broader syndrome, in which case the child may lose the benefit of early medical or surgical interventions. For example if a boy with unilateral renal agenesis is classified as having an isolated birth defect, but actually has Branchio-Oto-Renal (BOR) syndrome, he probably will not receive audiology screening (beyond the routine newborn screening) until his preschool health check. In this scenario, there may be delayed diagnosis of a hearing loss that is not present at birth, but that impacts speech and language acquisition.

Similarly if an apparently isolated cleft palate is actually part of microdeletion 22q11.2 syndrome, the child will not be evaluated for ectopically placed carotids and may suffer an adverse outcome during palate repair. In his textbook *Practical Genetic counseling*, Peter Harper states "*few cases ... (of microdeletion 22q11) ... are truly non-syndromal, but the features can be subtle*" (Harper, 2003⁴). Other baseline investigations that should be arranged once microdeletion 22q is diagnosed include:

calcium studies; absolute lymphocyte count and evaluation of the humoral immune system; and renal ultrasound. Furthermore, a person with microdeletion 22q11 should have a speech and language assessment by age 1 year because almost all affected individuals have delayed speech and benefit from early language intervention (McDonald-McGinn et al. 2005).

2. *Failure to arrange genetic testing.* The most common genetic tests indicated for children with a IIA diagnosis are chromosome analysis and FISH testing (to rule out a chromosomal syndromic etiology). These tests can be ordered by a non-geneticist as long as that physician recognizes that such testing is indicated. However there are instances where a specific molecular genetic test should be ordered, as described above.

Over the next decade, molecular genetic testing for isolated birth defects will almost certainly be incorporated into clinical practice. There are two categories of future clinically relevant genetic tests for birth defects:

- a) Tests for a panel of alleles (gene variants and or copy number variants) that have been identified as determining the genetic risk for a multifactorial malformation. Before such "DNA chips" can have clinical utility, the malformation's underlying genetic etiology must be fully understood, and this is likely to include multiple susceptibility and protective alleles.
- b) Tests for particular single-genes which when mutated produce a Mendelian form of an isolated birth defect.

At present only a handful of susceptibility genes for multifactorial congenital malformations have been identified. For example the *MTHFR* 677C->T allele was the first recognized genetic risk factor for folate-sensitive neural tube defects (NTDs). Fetal homozygosity for this allele is associated with up to a 7-fold increased risk of NTD in certain populations (Ou et al. 1996). Because any one such susceptibility gene confers only part of the genetic risk, these tests are not yet being offered clinically. Once the susceptibility and protective alleles for a particular multifactorial genetic disorder have been identified, probands and their parents can be genotyped for high and low risk markers. The algorithm for refining recurrence risk based on parental genotype will also need to include "gene x environment" interactions.

Genes that operate in a Mendelian way to produce isolated birth defects have been identified. Due to the rarity of such cases, these molecular tests are not in wide spread clinical use, although some examples were mentioned above. As the price of molecular genetic testing falls (Chung 2007), this situation is likely to change. Other examples of isolated malformations caused by single-gene mutations include:

- i. *CFC1*: In 2002, Goldmuntz et al. showed that autosomal dominant *CFC1* mutations cause not only heterotaxy syndromes, but in rare instances isolated complex congenital heart disease (transposition of the great arteries and double outlet right ventricle). Most patients have *de novo* mutations, but recurrence could occur if one parent had gonadal mosaicism for the mutation. A couple who have had a child with a *CFC1* mutation would have the option of prenatal molecular genetic testing in subsequent pregnancies.

- ii. Similarly several genes have been identified that cause isolated autosomal dominant cleft lip and/or palate including *MSX1* (Jezewski et al. 2003), *IRF6* (Kondo et al. 2002), *TP63* (Leoyklang et al. 2006) and *SUMO1* (Alkuraya et al. 2006). Only *IRF6* testing is clinically available (www.genetests.org). A patient with a bilateral cleft lip and palate who has an *IRF6* mutation has a 50% risk of having an affected child (rather than the multifactorial offspring recurrence risk of 5-6%).

For cleft lip and palate patients, pedigree review is important not only because it may identify an autosomal dominant family history of clefts, but also because it may reveal a cancer pattern suggestive of the presence of an *E-cadherin* mutation. Autosomal dominant mutations in this gene cause Hereditary Diffuse Gastric Cancer (HDGC) syndrome. The hallmark malignancies are diffuse gastric cancer and lobular breast cancer, but cleft lip with or without cleft palate is also a feature (Frebourg et al. 2006) and this is the only manifestation of the syndrome that would be present in a mutation-positive child. HDGC is over-represented in Newfoundland due to the presence of a *E-Cadherin* founder mutation (see section 2B.3)

3. *In the absence of a genetics consultation, the parents of a child with a IIA birth defect may not be counseled about their risk of having another affected child and may fail to be referred to a maternal fetal medicine specialist (high-risk obstetrician) during subsequent pregnancies.*

As an example, all children with holoprosencephaly (HPE) should have a chromosome analysis. 25-50% have a chromosome abnormality with specific recurrence risk figures depending on the abnormality. Another 5% have microdeletions which are usually *de novo* and associated with a low recurrence risk. These can be identified through genomic microarray analysis, and in most Canadian centres microarray testing can only be ordered by a clinical geneticist. If the above are both normal and the child appears to have a non-syndromic form of HPE, testing of the following autosomal dominant genes should be considered: *SHH*, *TGIF*, *SIX3* and *ZIC2*. These genes cause Mendelian non-syndromic forms of HPE and are associated with strikingly variable phenotypes even within families, ranging from alobar HPE to essentially normal individuals who have ocular hypotelorism (closely spaced eyes). If a mutation in one of these genes is identified and one of the parents is a carrier, the recurrence risk for subsequent pregnancies is 50% (Meunke and Gropman 2005).

5E.4 Referral rates for children with a IIIA diagnosis: The IIIA diagnoses are multifactorial diseases with high heritability, excluding birth defects. 778 children in the dataset had a IIIA diagnosis of whom 60 were referred (~8%).

Genetic referral is appropriate for a subset of children with a IIIA diagnosis (e.g. autism spectrum disorder, mental retardation), but in the absence of a Mendelian looking family history, referral is not indicated for most of these conditions (e.g. diabetes, asthma, inflammatory bowel disease). There is also a group of IIIA diagnoses for which referral to genetics *may* be appropriate (ADHD, behavioral abnormalities, seizures, some psychiatric conditions including schizophrenia).

In our dataset, the two most common IIIA diagnoses for which genetic consultation is indicated are autism spectrum disorder (ASD) (Chudley et al. 1998; Battaglia and Carey 2006; Schaefer and Mendelson 2008) and mental retardation/developmental delay (Fryns et al. 1986).

5E.4.1 Referral rates for children with a IIIA diagnosis – Autism: *Within the IIIA group, there were 16 autistic children of whom 10 were referred (64%).* Every autistic child should ideally be assessed by a medical geneticist; about 25% have significant dysmorphology and in up to half of these, the physical findings lead to a syndromic diagnosis (Miles et al. 2005). Each affected child should have routine chromosome analysis and DNA testing for Fragile X syndrome (*FMR1* gene), with diagnostic yields of 5-10% and 5% respectively (Miles and McCathern 2005; Reddy 2005; Schaefer and Mendelsohn 2008). If the child tests positive for an *FMR1* mutation, his or her parents have a 25% risk of having a son with Fragile X syndrome and prenatal diagnosis is available.

Admittedly the following information about the genetic etiology of autism was not known during the 2000-2001 period over which the study children were hospitalized, however our improved understanding of the genetics of ASD underscores the importance of a clinical genetics consultation:

- i. Firstly, recent publications recommend that genomic microarray analysis be considered for any autistic child if the above routine genetic blood work is normal (Marshall et al. 2008). The microarray screens the genome for dosage imbalances and

is one method of identifying copy number variants (CNVs) at the 16p11.2 locus. Microdeletions and microduplications at this locus occur in 1% of autistic children (Marshall et al. 2008; Weiss et al. 2008). This CNV is usually a *de novo* change so the parents of a child with a 16p11.2 CNV have a recurrence risk that is less than the 5-10% empiric sibling recurrence risk for “multifactorial” autism. This lower recurrence risk may influence the parents’ decision about having more children.

- ii. Secondly, several Mendelian forms of ASD have been recognized and genetic testing for these is indicated in selected cases. Examples include *neuroligin-3* and *neuroligin-4* testing in patients with an X-linked family history and *PTEN* testing in autistic children with macrocephaly greater than 4 standard deviations above the mean. (Miles and McCathern 2005; Jamain et al. 2003; Laumonnier et al. 2004; Butler et al. 2005). Any individual with a *PTEN* mutation is believed to be at increased risk for developing Cowden syndrome related malignancies and cancer screening is recommended beginning at age 18 (Zbuk et al. 2006).

Schaefer and Mendelsohn (2008) made the strongest published recommendation to date for referring autistic patients to a clinical geneticist. The authors propose a three tiered clinical genetics evaluation and suggest that it will have an overall diagnostic yield of 40%. The first tier includes: dysmorphology and Woods lamp examinations (the latter for pigmentary abnormalities of the skin); chromosome analysis; and DNA testing for Fragile X syndrome. The second tier includes genomic microarray, with the following additional tests targeted to specific patient subsets: skin karyotype for patients with significant pigmentary abnormalities; sequencing of the *MECP2* gene for all female ASD

patients; and *PTEN* testing for individuals with a head circumference $>+2.5SD$. Tier 3 investigations include brain MRI and methylation analysis of chromosome 15.

The authors emphasize that recent advances have created the opportunity for clinical geneticists to move from empiric recurrence risk counseling to diagnosis-related counseling for patients with autism.

5E.4.2 Referral rates for children with a IIIA diagnosis – mental retardation (MR):

Within the IIIA group, there were 59 children with mental retardation/developmental delay of whom 18 were referred (31%).

Bodensteiner and Schafer (1995) state that the *"association of mental retardation and congenital malformations has long been recognized"* and that *"a necessary component of the evaluation of the child with idiopathic mental retardation is a comprehensive dysmorphologic examination."* The literature suggests up to 50% of individuals with MR have an identifiable etiology. Once environmental causes have been excluded, the dysmorphology examination is used to categorize the patient as having syndromic or non-syndromic MR. Any child with delayed speech, language or motor development of unknown etiology should have a chromosome analysis and DNA testing for Fragile X syndrome. Other targeted investigations include metabolic testing, brain MRI, genomic microarray analysis and confirmation of suspected specific syndromic forms of MR by molecular genetic testing (Firth and Hurst 2005; Curry et al. 1997). In the future, DNA chips will probably be used to test multiple MR genes simultaneously.

5E.4.3 Referral rates for children with other IIIA diagnoses: Within the unique dataset of 778 children with a IIIA diagnosis, there were some specific diagnoses for which genetic referral may be indicated depending on the presence of other clinical features and/or a positive family history. The referral patterns for some of these include:

- i. ADHD and/or behavioral abnormalities: 10/52 (19.2%) were referred.
- ii. Non-febrile seizures: 8/82 were referred (9.8%). There were two cases of infantile spasms and both were not referred. Infantile spasms develop in up to 1/3 of children with tuberous sclerosis which is an autosomal dominant disease (Curatolo et al. 2008).
- iii. Psychiatric disorders: 3/78 referred. The diagnoses of the three patients that were referred were: depression, bipolar disorder with autism, and obsessive compulsive disorder with autism.

5E.5 Deficits in genetic consultations vs. existing clinical genetic resources: During a 14-month period, 3,281 unique hospitalizations generated 365 genetic consultations (~9% of all the hospitalizations). Although data on the timing and location of these consultations was not extracted, it is reasonable to assume that these occurred as a mixture of inpatient and outpatient assessments, and that these consults could have occurred either before, during or after the index hospitalization. The dataset contains another 129 children who had a diagnosis that was an unequivocal indication for genetic referral, but who were not referred. This number was calculated assuming that all IA-IE children and that 60% of IIA children should have been referred. Children with a IIIA

diagnosis of either developmental delay or autism were also included as shown in the following table (*table 5-5*).

Table 5-5: *Children with a final genetic category that was an indication for genetic consultation who were not referred for genetic services:*

Category (Most genetic final diagnosis)	Number of children NOT referred
IA	10
IB	30
IC	6
ID	2
IE	20
IIA	61
Selected IIIA	41 DD + 6 ASD
TOTAL	129

Assuming that none of these 129 children had a genetic consult outside the province that was not evident in the discharge summary, the total number of referred cases should have been $365 + 129$, i.e. $494/3281$ or 15.1% of the unique dataset. Put another way, for almost every three children that were referred, there was one child who was inappropriately not referred.

Taking this one step further, assuming that 1999/2000 referral pattern is similar to the present one, there is a genetic referral deficit of 129 children over an 14-month period or 111 children annually.

Although the PMGP is part of *Eastern Health*, it has a provincial mandate to deliver genetic services. Present staff includes two clinical geneticists and eight genetic counselors (GC's), including two GC's that work in outreach clinics in Gander and Corner Brook. The physicians see all the referred patients who do not yet have a specific genetic diagnosis; many of these individuals require a physical examination. The genetic counselors have relatively independent caseloads of patients with known genetic

diagnoses who require genetic counseling and/or genetic testing. Almost all of the “inappropriately not referred” children described above would have required a physician rather than genetic counselor assessment.

During the 2005-2006 fiscal year, 1,621 patients were referred to the PMGP and 1,805 patients were seen by a clinical geneticist and/or a genetic counselor. The distribution of the types of consults was: 15% prenatal; 40% pediatric; and 45% non-prenatal adult cases. 88 of the 1,805 consultations were inpatient consults which were done by one of the clinical geneticists, and over 90% of these were neonatal or pediatric cases. The two clinical geneticists each saw ~400 patients (i.e. a total of 800 patients including the 88 inpatient consults). The remaining 1,000 cases were seen by a genetic counselor and supervised and then co-signed by one of the two clinical geneticists.

The present wait time for semi-urgent and non-urgent patients to see a clinical geneticist is six months and two years respectively. In order to accommodate the non-referred children without increasing the present wait time, each of the clinical geneticists would have to assess ~12.5% more patients per year (an additional ~50 patients annually, representing 1.5 extra month's worth of patients).

5E.6 Future of genetic testing for complex diseases and the role of medical

geneticists/genetic counselors: The Human Genome Project was a 13-year international effort that was coordinated by the US Department of Energy and the National Institute of Health, and one of its major goals was to sequence the entire human genome. This was achieved in 2003 when 95% of the gene containing part of the human genome was

finished to 99.99% accuracy (Collins et al. 2003). The genome is estimated to contain 25,000 genes over 2,100 of which are associated with a monogenic human disease. Clinical genetic tests are available for >1100 disorders (www.Genetests.org). These phenotypes are within the traditional realm of clinical geneticists and genetic counselors.

Completion of the human sequence, coupled with sequencing of the genomes of several important model organisms (including yeast and drosophila) and characterization of the human haplotype structure, provided the tools necessary for large scale genome-wide association studies. These have started to identify genes that harbor variations which are associated with common multifactorial diseases. Examples that are relevant to the pediatric population include type I diabetes, obesity (Frayling et al. 2007), inflammatory bowel disease (Duerr et al. 2006; Duerr 2007) and autism (Losh et al. 2008).

Another previously undetected source of variation in the genome is differences in the copy number of genes. Copy number variants (CNVs) are detected by genome microarrays and appear to underlie a significant fraction of birth defects, mental retardation and autism. It has been hypothesized that CNVs also increase susceptibility to other more common psychiatric and medical conditions (Fernandez et al. 2009; Cook and Scherer 2008; Shaffer and Bejjani 2006; Zahir and Friedman 2007).

Numerous reviews and editorial articles have been published which discuss the fact that we are about to enter the era of “molecular medicine” (Chung 2007; Woodcock 2007; Sieberts and Schadt 2007). Individual genetic and genomic information will almost certainly be integrated into health care provision and will be increasingly utilized in every

field of medicine. Treatment will become targeted less by symptoms and more by understanding the fundamental causes of disease, and it has been predicted that future drug design will be based on knowledge of perturbed genes.

The potential benefits of molecular medicine include:

1. Refining the treatment of common diseases. An individual's genetic variations may be used to select a particular drug and/or to determine dosing (see *section 5E.7* for further discussion).
2. Preventing disease by identifying genetic variants that make certain individuals susceptible to an adverse outcome from a particular environmental factor (e.g. lung cancer from smoking). Similarly, individuals with particular genotypes may benefit from a preventive medical therapy or increased surveillance. This may occur through population screening.

That being said, it is clear that the genetic basis for most common diseases is complex involving multiple genetic variants (gene polymorphisms and possibly copy number variants) and interactions with the environment. Algorithms for predicting the development of common diseases will not come into clinical use until all the genetic determinants of a disease have been identified (both susceptibility and protective variants). Also "gene x gene" and "gene x environment" interactions must be factored in to the composite risk calculation. To date, there is no complex disease for which the underlying genetics is this completely understood.

Currently, genetic medicine is delivered by physician specialist medical/clinical geneticists and masters trained genetic counselors. In Canada, there are 106 practicing

medical geneticists who have been certified through the Royal College of Physicians and Surgeons and/or the Canadian College of Medical Geneticists (*personal communication, Roberta Sulpha, CCMG secretariat, 2008*). There are ~250 genetic counselors certified through the Canadian Association of Genetic Counsellors (CAGC) who are employed in Canada.

There is agreement in the literature and within the Canadian clinical genetics community that once genetic testing for common complex diseases becomes part of routine medical practice, the volume of such genetic testing will rapidly overwhelm traditional genetic clinics. These tests will almost certainly be ordered by family physicians and non-geneticist specialists who will use the tests to make clinical decisions. Some individuals who have had genetic testing for complex diseases may be flagged for referral for genetic counseling. The genetic counselor who provides this may work in a traditional genetics clinic. Alternatively, the counselor may work in a subspecialty clinic or may be under the supervision of a diagnostic laboratory director.

Assuming complex disease testing becomes part of the general practice of medicine, the following is a breakdown of children from our dataset who would require genetic services (*table 5.6*).

We have assumed that no child within the IIB, IIIB, IV or V group would require genetic services (clearly not the case because some of these children were referred). The calculation used the simplified assumption that all IA-IE children require genetic consultation that would be provided by a clinical geneticist working with the assistance of a genetic counselor. We assumed that 60% of the IIA group would require the same,

with the rest of the group being eligible for complex trait genetic testing that would be organized outside a traditional genetics clinic. Most of the IIIA group (with diagnoses that include Type I diabetes and other autoimmune diseases, most psychiatric disorders and inflammatory bowel disease) would have the option of complex disease genetic testing ordered by a non-geneticist physician, and only the MR/ developmental delay and autism groups would receive traditional genetic services through a clinical geneticist or genetic counselor.

Table 5-6: Breakdown of patients from the unique admissions dataset (3,281 patients) who will require genetic services once complex disease genetic testing is routinely available.

Hierarchical Final Category of hospitalization	Number of patients	Number to be seen by a clinical geneticist	Number to receive genetic services outside genetics clinic
IA	30	30	-
IB	120	120	-
IC	7	7	-
ID	10	10	-
IE	71	71	-
IIA	217	130	87
IIB	137	-	-
IIIA	778	*75	703
IIIB	809	-	-
IV	1,063	-	-
V	39	-	-
TOTAL	3,281	443	790

* there were 16 unique patients with autism and 59 with mental retardation or developmental delay

The above table shows that within the dataset of 3,281 unique hospitalizations over a 14-month period, 13.5% presently require genetic services from a clinical geneticist and another 24% are likely to eventually require complex disease genetic testing through a non-geneticist physician.

The model has not addressed pharmacogenomic testing which may be relevant to many of the category IV children.

5E.7 Pharmacogenomics: Pharmacogenomics refers to the concept of screening for genetic variants that determine drug metabolism and adverse drug effects. Differences in drug metabolism have been linked to genes that encode drug metabolism enzymes, drug transporters and drug targets. The goal of pharmacogenomics is to avoid the trial and error approach to drug therapy by screening patients for such variants. The patient's genetic profile will be used to select the "best drug" at the "best dose". This approach will improve the efficacy of prescribed medications and reduce side effects, which should in turn increase patient compliance. At present between 20-40% of patients fail to respond to commonly used medications for diabetes, hypertension and depression (Haga and Burk 2004).

Unlike genetic testing for multifactorial diseases, which presently has limited clinical utility, some pharmacogenomic tests are being ordered by physicians for patient management. As an example, the enzyme thiopurine methyltransferase (TMPT) inactivates thiopurine drugs like azathioprine, thioguanine and 6-mercaptopurine. These drugs are used to treat patients with conditions that include inflammatory bowel disease, rheumatoid arthritis, acute lymphoblastic leukemia and organ transplant rejection. One in 300 people is homozygous for a variant that encodes a form of the enzyme with reduced activity. When such a person is given a thiopurine drug, the active compound accumulates leading to potentially fatal bone marrow failure. Some physicians now screen for this variant prior to administration of a thiopurine drug and administer a lower dose if the patient carries two copies of the variant. It is worth noting that any siblings of such a patient are at 25% risk of having the same genotype and should be counseled accordingly (Zhou 2006; Firooz et al. 2008; Tsui 2008).

Cytochrome p450 is a complex of heme containing enzymes. Two of these genes, *CYP2D6* and *CYP2C19* are responsible for the metabolism of 25% of drugs. Roche diagnostics has marketed the AmpliChip *CYP450* test. The chip assays 31 variants in these two genes and the results are integrated to produce a simple interpretation – the patient is predicted to be a poor, intermediate, extensive or ultrarapid metabolizer (Chung 2007). As the field of pharmacogenomics moves forward, it is easy to imagine that a substantial proportion of patients who are prescribed a medication might have a genetic test so that they can be dosed according to their drug metabolizing phenotype. Hence many of the category IV children (*table 5-5*) might also have been candidates for molecular genetic testing if screening for variation in drug metabolism was part of the routine practice of medicine.

5F. STRENGTHS AND LIMITATIONS OF THE STUDY:

5F.1 Strengths: One of the strengths of this study is the large number of hospitalizations that were reviewed, 4,144 consecutive separations over a 14-month period to the only tertiary-level pediatric hospital in Newfoundland and Labrador. The method of determining the genetic content of each hospitalization (review by a research nurse of the ICD-9 discharge codes, of the index discharge summary and of up to two previous discharge summaries if these existed) was validated against a gold standard (review of patient and medical record by a clinical geneticist) and was shown to be highly accurate. None of the previous pediatric admission studies included a validation component. The sample size was comparable to the Hall (1978) and McCandless (2004) publications and gives excellent statistical strength to the findings.

While all the previous pediatric hospitalizations studies clearly identified children with strongly genetic conditions (chromosomal and Mendelian diseases), our classification scheme of multifactorial conditions was unique and we believe more meaningful than the various earlier approaches. We divided multifactorial conditions into birth defects versus non-congenital disorders, and further subdivided each into high and low genetic content groups based on heritability and/or on the sibling recurrence risk (λ_s) for each disorder.

Another unique feature of our study is that the sample was drawn from a founder population. Because Newfoundland has been recognized as one of a handful of founder populations, the amount and distribution of genetic disease in this province may be different than in more admixed populations. This is the first Newfoundland study that has attempted to quantitate the burden of genetic disease using, as a sample of convenience, inpatients from the only pediatric hospital in the province.

Our analysis of hospital utilization data was more comprehensive than any of the previous studies, most of which were limited to mean length of stay (we also examined cumulative hospital days and cumulative surgeries).

Finally apart from Scriver et al. 1973, ours is the only study that examined the clinical genetic referral rates for children with strongly genetic diagnoses (and with particular moderately genetic diagnoses).

5F.2 Limitations: The validation study was performed using only one medical geneticist, with no inter-observer reliability estimate.

In the main study, we examined the burden of genetic disease among hospitalized children primarily by determining the *prevalence* of particular categories of disease. We examined disease *severity* only indirectly, by looking at differences in utilization of hospital resources between different genetic content groups.

Our ability to classify the genetic content of each admission was limited by the conditions that were recognized and written down by the admitting health care team. Although the method of categorizing the separations was validated, it was retrospective. Also we somewhat arbitrarily divided multifactorial diseases into two groups based on either sibling recurrence risk (with a cut-off of 10) or heritability estimate (with a cut-off of 50%).

Our ability to accurately classify the hospitalizations was also limited by our current understanding of the genetic basis of the disorders. In particular, it may turn out that a proportion of certain diseases that are presumed to be multifactorial are in fact strongly genetically determined. As an example, while many children with non-syndromic mental retardation have multifactorial MR, some have single-gene forms or pathogenic copy number variants (CNVs). Identification of the former group is still hampered by the cost of genetic testing. Microarray testing identifies the latter. While microarrays are now routinely ordered by clinical geneticists, none of the children in the dataset would have had this test. Among the MR population with a normal routine karyotype, genomic microarray testing identifies a pathogenic genomic deletion or duplication in 10-15%

(Stankiewicz and Beaudet, 2007). In this study, children with MR were classified as having a group 2 or moderately genetic disorder. If all of these had been referred for genetic consultation and if microarray testing had been in routine clinical use (as it now is), up to 15% of these children would have been categorized to the strongly genetic group (group 1).

Another limitation is that the study is not truly population-based. While 60% of Newfoundland children who need hospitalization are admitted to the Janeway hospital, the rest are admitted to community hospitals throughout the province. Because many of these hospitals do not have computerized discharge summaries, these children were excluded from the study. Because the Janeway hospital is the only pediatric hospital in the province, some referral bias also exists. Sicker and more medically complicated children are more likely to be admitted to the Janeway hospital even if there is a local community hospital that accepts children.

Also this study did not include neonates either born in the hospital or admitted to the neonatal intensive care unit. The objective of the study was to examine the genetic disease present among patients admitted to hospital, so that newborns born in the hospital and admitted to the nursery as part of routine care were excluded. We agree with McCandless et al. (2004) that the neonatal intensive care unit population warrants its own study, with classification criteria specifically designed for them. This population is likely to contain a higher proportion of patients with strongly genetic disorders and birth defects. Some of these probably have such severe phenotypes that they do not survive the neonatal period and hence would not be seen in a pediatric inpatient sample.

5G CONCLUSIONS: Newfoundland has been recognized as one of the world's relatively few "very young" founder populations. Due to genetic drift and founder effect, certain Mendelian diseases are more common in this province than in outbred populations, with many of the documented examples being adult-onset disorders. This is the first study that attempted to quantitate the net burden of genetic disease in Newfoundland, using as a sample of convenience 4,144 consecutive pediatric admissions to the only tertiary-level children's hospital in this province.

We validated our study methods (retrospective determination of the genetic content of an admission through review of discharge codes and discharge summaries) against a gold standard (review of the patient and hospital chart by a medical geneticist). We found that the retrospective review had a high level of accuracy, presumably because each separation's discharge summary was carefully reviewed by a single research nurse rather than relying on discharge codes alone to categorize the separations.

Our dataset consisted of 4,144 consecutive children (neonates were excluded) who were hospitalized over a 14-month period in 2000 and 2001. We found that 7.32% were admitted because of a Mendelian or probable Mendelian genetic syndrome (303 children) and that ~1% were admitted because of a chromosomal disorder (39 children). Hence diseases that are entirely (or nearly entirely) genetically determined were responsible for 8.3% of the consecutive hospitalizations. Children with higher heritability multifactorial diseases (birth defects and diseases that are not congenital) accounted for ~25% of the dataset (1,033 admissions). The remaining two-thirds of the admissions occurred because of a minimally or non-genetic condition.

There have been five previous studies of the burden of genetic disease among children admitted to a single pediatric hospital. In order to make our data more comparable to that of the earlier studies, each separation's diagnosis list was reanalyzed and the hospitalization was assigned to a genetic content category (the hierarchical final category) based on the diagnosis that had the highest genetic ranking, even if it did not contribute to the child's need to be hospitalized or influence the hospital stay. 8.4% of the admitted children had a strongly genetic diagnosis (7.4% single-gene and 1% chromosomal) and 30.6% had a moderately genetic disease. The frequency range for Mendelian and chromosomal diseases in the earlier studies was 4-9.5% and 0-1.29% respectively (*table 5-2*). Hence our initial hypothesis was not correct; the proportion of hospitalized NL children with strongly genetic diseases was *not* higher than previously documented in admixed populations. Most of the founder effect mutations that have been identified in the Newfoundland population are adult-onset diseases (HNPCC, ARVD5, MEN1 etc). Hence while there may be some over-represented Mendelian diseases that affect children (a known example is infantile neuronal ceroid lipofuscinosis), when aggregated as a group and compared to previous single hospital studies, single-gene disorders do not appear to disproportionately burdening Newfoundland's pediatric inpatient services.

Moreover our dataset contained a lower proportion of Newfoundland children with apparently multifactorial birth defects than any of the five earlier studies (*table 5-2*). None of the previous studies divided congenital anomalies into two groups based on sibling recurrence risk, so we can compare only the *total* birth defect frequencies. In our study, 11.2 % of the children had either a high or a low heritability birth defect. Only

three of the five earlier studies used a classification scheme that determined the frequency of children with non-syndromic birth defects which ranged from 18.5% - 22.4% (Scriver et al. 1973; Hall et al. 1978; Carnevale et al. 1985).

Compared with earlier publications, conditions that were more common in our consecutive admission dataset were the group of higher heritability multifactorial diseases (excluding birth defects). These diseases include asthma, autism, type I diabetes, inflammatory bowel disease and juvenile rheumatoid arthritis (JRA). These IIIA multifactorial diseases were present in almost one-quarter of our consecutive admissions (23.4%), compared with frequencies of 8.5-10.7% in three of the earlier studies (Day and Holmes 1973; Hall et al. 1978; Carnevale et al. 1985). Unfortunately the comparable proportion from McCandless' 2004 study, which is the most recent of the earlier studies, could not be determined. While some of these diseases have become more prevalent worldwide over the past 30 years, it is possible that a subset (for example type I diabetes) is over-represented in our province because of the population's genetic architecture, i.e. because of drift of disease-associated alleles in the Newfoundland population compared with the parental Irish and English populations.

Our study sample included 120 unique children who were hospitalized with a single-gene disorder, and roughly equal numbers had an autosomal dominant disease versus an autosomal recessive one (40.2% and 39.4% respectively). The proportion of NL children with autosomal recessive disorders was lower than in Hall's study (1978) where 56.8% of the children with Mendelian disorders had an autosomal recessive disease. This finding is consistent with the hypothesis that much of the burden of genetic disease in this province

is due to autosomal dominant disorders which have become prevalent due to large numbers of families with high sibship sizes, rather than due to autosomal recessive disorders which collectively are still fairly rare. Put another way, even in founder populations like Newfoundland, the vast majority of mutant autosomal recessive alleles exist in healthy carriers rather than in affected individuals.

We have shown that children with strongly genetic diseases collectively use more hospital resources than those with less genetic conditions. Children with strongly genetic diseases (group1) had an average length of stay that was twice as long as those admitted with minimally genetic conditions (8.01 vs. 3.99 days). This was consistent with the McCandless et al. 2004 study. Moreover, we determined that children in the strongly genetic group had 1.8-fold more surgeries when corrected for age, compared with those in the moderately and minimally genetic groups. Finally when corrected for age, group 1 children had 3.8-fold more cumulative hospital days than those in the minimally genetic group.

Because of the strength of our length of stay (LOS) data, we used multivariate regression to examine the impact of the genetic content of an admission on LOS. We ran three regression analyses, but the genetic content of the admission explained no more than 4.5% of the variance in LOS, compared with 34% of the variance which was explained by the admitting diagnosis. Nevertheless, we determined that children with teratogenic exposures (IC diagnoses) and those with single-gene disorders (IB) were at the highest risk for prolonged hospitalization (> 7 days).

When we analyzed referral rates for genetic consultation, the largest referral deficits occurred for the high heritability birth defects (only one-third referred) and for certain high heritability complex diseases including autism and unexplained mental retardation (again only one-third referred). During the 14-month study period, 3,281 unique hospitalizations generated 365 genetic consultations and we calculated that an additional 129 children should have been referred. Fifteen percent of children in the unique data set had a diagnosis which is presently an indication for genetic consultation. The existing provincial medical genetics service would have difficulty coping with these additional assessments (about 110 extra children annually). Moreover, we estimated that if these same 3,281 children were to be hospitalized ten years from now, an *additional* 24% would require complex disease genetic testing which will probably be ordered by physicians who are not geneticists.

In summary, we have determined that children with strongly and moderately genetic diseases are responsible for about 1/3 of pediatric admissions. The proportion of children with chromosomal and Mendelian diseases was comparable to earlier studies from more admixed populations, and the proportion of hospitalized Newfoundland children with birth defects was in fact lower. Compared to these earlier studies, children with multifactorial diseases (excluding birth defects) accounted for a larger fraction of our separations, and this may reflect increased disease prevalence related to Newfoundland's unique genetic architecture and/or to particular environmental factors acting on a genetically vulnerable background. Finally, we have shown that children with wholly genetic diseases (who account for about 8.3% of all admissions) use disproportionately

large amounts of inpatient hospital resources, and this information should be incorporated into future healthcare work force planning.

REFERENCES:

- Abramsky L, Botting B, Chapple J, Stone D (1999): Has advice on periconceptional folate supplementation reduced neural-tube defects? *Lancet* 354:998-999.
- Allderdice PW, Browne N, Murphy DP (1975): Chromosome 3 duplication q21 leads to qter deletion p25 leads to pter syndrome in children of carriers of a pericentric inversion inv(3) (p25q21). *Am J Hum Genet* 27:699-718
- Alkuraya FS, Saadi I, Lund JL, Turbe-Doan A, Morton CC, Maas RL (2006): SUMO1 haploinsufficiency leads to cleft lip and palate. *Science* 313:17
- American Academy of Pediatrics, Committee on Genetics (2001): Health supervision for children with Down syndrome. *Pediatrics* 107:442-9.
- Baird PA, Anderson TW, Newcombe H, Lowry RB (1988): Genetic Disorders in Children and Young adults: A population study. *Am. J Hum Genet* 42:677-693.
- Battaglia A, Carey JC (2006): Etiologic yield of autistic spectrum disorders: a prospective study. *Am J Med Genet C Semin Med Genet* 142:3-7.
- Bear JC, Briones-Urbina R, Fahey JF, Farid NR (1985): Variant multiple endocrine neoplasia I (MENI Burin): further studies and non-linkage to HLA. *Hum Hered* 35:15-20.
- Bear JC, Nemec TF, Kennedy JC, Marshall WH, Power AA, Kolonel VM, Burke GB (1987): Persistent genetic isolation in outport Newfoundland. *Am J Med Genet* 27:807-830.
- Bear JC, Nemec TF, Kennedy JC, Marshall WH, Power AA, Kolonel VM, Burke GB (1988): Inbreeding in outport Newfoundland. *Am J Med Genet* 29:649-60.
- Beresford RG, Tatlidil C, Riddell DC, Welch JP, Ludman MD, Neumann PE, Greer WL (2000): Absence of fragile X syndrome in Nova Scotia. *J Med Genet* 37:77-9
- Bodensteiner JB, Schaefer GB (1995): Evaluation of the patient with idiopathic mental retardation. *J Neuropsychiatry Clin Neurosci* 7:361-70.
- Bowen P. Workshop on genetic disorders in the Hutterites--Edmonton, Canada, October 12-13, 1983 (1985): *Am J Med Genet* 22:449-51.
- Butler MG, Dasouki MJ, Zhou XP, Talebizadeh Z, Brown M, Takahashi TN, Miles JH, Wang CH, Stratton R, Pilarski R, Eng C (2005): Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *J Med Genet* 42:318-21.

Carnevale A, Hernandez M, Reyer R, Paz F, Sosa C (1985): The frequency and economic burden of genetic disease in a pediatric hospital in Mexico City. *Am J Med Genet* 20:665-675.

Carter CO (1965): The inheritance of common congenital malformations. *Progr Med Genet* 4:59.

Cavanagh J, Mathews M, Crane J (2007): Awareness and use of maternal serum screening among women from the St. John's region of Newfoundland and Labrador. *J Obstet Gynaecol Can* 29:630-4.

Chudley AE, Conry J, Cook JL, Loock C, Rosales T, LeBlanc N; Public Health Agency of Canada's National Advisory Committee on Fetal Alcohol Spectrum Disorder (2005): Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. *CMAJ* 172(5 Suppl): S1-S21.

Chudley AE, Kilgour AR, Cranston M, Edwards M (1998): Outcomes of genetic evaluation in children with pervasive developmental disorder. *J Dev Behav Pediatr* 19: 321-5

Chung WK (2007): Implementation of genetics to personalize medicine. *Gend Med* 4: 248-65.

Citizens IS (1977): *The History of Burin*. Marystown, Canada, South Coast Printers.

van Cleve SN, Cohen WI (2006): Part I: clinical practice guidelines for children with Down syndrome from birth to 12 years. *J Pediatr Health Care* 20:47-54.

Collins FS, Morgan M, Patrinos A (2003): The Human Genome Project: lessons from large-scale biology. *Science* 300:286-90.

Cook EH Jr, Scherer SW (2008): Copy-number variations associated with neuropsychiatric conditions. *Nature* 455:919-23.

Cunniff C, Carmack JL, Kirby RS, Fiser DH (1995): Contribution of heritable disorders to mortality in the pediatric intensive care unit. *Pediatrics* 95:678-81.

Curatolo P, D'Argenzio L, Cerminara C, Bombardieri R (2008): Management of epilepsy in tuberous sclerosis complex. *Expert Rev Neurother* 8:457-67.

Curry CJ, Stevenson RE, Aughton D, Byrne J, Carey JC, Cassidy S, Cunniff C, Graham JM Jr, Jones MC, Kaback MM, Moeschler J, Schaefer GB, Schwartz S, Tarleton J, Opitz J (1997): Evaluation of mental retardation: recommendations of a Consensus Conference: American College of Medical Genetics. *Am J Med Genet* 72: 468-77.

Czeizel A, Dudas I (1992): Prevention of the first occurrence of neural tube defects by periconceptional vitamin supplementation. *N Eng J Med* 327:1832-1835.

Day N, Holmes L.B. (1973): The incidence of Genetic Disease in a University Hospital Population. *Am J Hum Genet* 25:237-246.

Dohey A, Crowley M, Green R, Younghusband B, Fernandez B: Is FRAXE over represented among Newfoundland boys with cognitive impairment? Abstract *European Society of Human Genetics*, May 2008

Dubois-Laforgue, D, et al. (1997). Insulin-dependent diabetes mellitus in non-DR3/non-DR4 subjects. *Hum Immunol* 57:104-109.

Duerr RH (2007): Genome-wide association studies herald a new era of rapid discoveries in inflammatory bowel disease research. *Gastroenterology* 132: 2045-9.

Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barnada MM, Rotter JI, Nicolae DL, Cho JH (2006): A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314:1461-3.

Escamilla MA, Spesny M, Reus VI, Gallegos A, Meza L, Molina J, Sandkuijl LA, Fournier E, Leon PE, Smith LB, Freimer NB (1996): Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. *Am J Med Genet* 67:244-53.

Fernandez BA, Roberts W, Chung B, Weksberg R, Meyn S, Szatmari P, Joseph-George AM, MacKay S, Whitten K, Noble B, Vardy C, Crosbie V, Luscombe S, Tucker E, Turner L, Marshall CR, Scherer SW (2009): Phenotypic spectrum associated with *de novo* and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. *J Med Genet* [Epub ahead of print Sept 15, 2009]

Firooz A, Ghandi N, Hallaji Z, Chams-Davatchi C, Valikhani M, Karbakhsh Davari M (2008): Role of thiopurine methyltransferase activity in the safety and efficacy of azathioprine in the treatment of pemphigus vulgaris. *Arch Dermatol* 144:1143-7.

Firth HV, Hurst JA 9 (2005): Clinical Approach – Mental retardation, Ch 2 in: *Oxford Desk Reference: Clinical Genetics*. 1 ed. Oxford University Press. pp168-171.

Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI (2007): A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316:889-94.

Frebourg T, Oliveira C, Hochain P, Karam R, Manouvrier S, Graziadio C, Vekemans M, Hartmann A, Baert-Desurmont S, Alexandre C, Lejeune Dumoulin S, Marroni C et al. (2006): Cleft lip/palate and CDH1/E-cadherin mutation in families with hereditary diffuse gastric cancer. *J Med Genet* 43:138-142.

Friedman JM, Hanson JW (2002): Clinical teratology, Ch 39 in *Emery and Rimoin's Principles and Practice of Medical Genetics*. 4th ed. Harcourt Publishers Limited.

Friel JK, Frecker M, Fraser FC (1995): Nutritional patterns of mothers of children with neural tube defects in Newfoundland. *Am J Med Genet* 55:195-9.

Froggatt NJ, Green J, Brassett C, Evans DG, Bishop DT, Kolodner R, Maher ER (1999): A common MSH2 mutation in English and North American HNPCC families: origin, phenotypic expression, and sex specific differences in colorectal cancer. *J Med Genet* 36:97-102.

Fryns JP, Kleczkowska A, Dereymaeker A, Hoefnagels M, Heremans G, Marien J, van den Berghe H (1986): A genetic-diagnostic survey in an institutionalized population of 173 severely mentally retarded patients. *Clin Genet* 30:315-23.

Gardner RJM, Sutherland GR (1996): *Chromosome Abnormalities and Genetic Counseling*; 2nd ed. Oxford University Press.

Goldmuntz E, Bamford R, Karkera JD, dela Cruz J, Roessler E, Muenke M (2002): CFC1 mutations in patients with transposition of the great arteries and double-outlet right ventricle. *Am J Hum Genet* 70:776-80.

Green JS, O'Driscoll M, Barnes A, Maher ER, Bridge P, Shields K, Parfrey PS. (2002) Impact of gender and parent of origin on the phenotypic expression of hereditary nonpolyposis colorectal cancer in a large Newfoundland kindred with a common MSH2 mutation. *Dis Colon Rectum* 45:1223-32.

Green JS, Parfrey PS, Harnett JD, Farid NR, Cramer BC, Johnson G, Heath O, McManamon PJ, O'Leary E, Pryse-Phillips W (1989): The cardinal manifestations of Bardet-Biedl syndrome, a form of Laurence-Moon-Biedl syndrome. *NEJM* 321:1002-1009.

Haga SB, Burke W (2004): Using pharmacogenetics to improve drug safety and efficacy. *JAMA* 291:2869-71.

Hall JG, Powers EK, McIlvaine RT, and Ean VH (1978): The frequency and financial burden of genetic disease in a pediatric hospital. *Am J Med Genet* 1:417-438.

Hao W, Skarulis MC, Simonds WF, Weinstein LS, Agarwal SK, Mateo C, James-Newton L, Hobbs GR, Gibril F, Jensen RT, Marx SJ (2004): Multiple endocrine neoplasia type 1 variant with frequent prolactinoma and rare gastrinoma. *J Clin Endocrinol Metab* 89:3776-84.

**This page is
missing from the
original book**

Jones KL (2006): *Smith's Recognizable Patterns of Human Malformation*, 6 ed, Elsevier Saunders.

Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, Stengård J, Kesäniemi YA (1992): Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 35:1060-7.

Karvonen M, Viik-Kajander M, Moltchanova E, Libman I, LaPorte R, Tuomilehto J (2000): Incidence of childhood type I diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group. *Diabetes Care* 23:1516-1526.

Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N, Suriano G, Zaor S, Van Manen L, Gilpin C, Nikkel S, Connolly-Wilson M, Weissman S, Rubinstein WS, Sebold C, Greenstein R, Stroop J, Yim D, Panzini B, McKinnon W, Greenblatt M, Wirtzfeld D, Fontaine D, Coit D, Yoon S, Chung D, Lauwers G, Pizzuti A, Vaccaro C, Redal MA, Oliveira C, Tischkowitz M, Olschwang S, Gallinger S, Lynch H, Green J, Ford J, Pharoah P, Fernandez B, Huntsman D (2007): Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. *JAMA* 297:2360-72.

King RA, Rotter RA, Motulsky G (2002): *The Genetic Basis of Common Diseases*. 2nd ed. Oxford University Press.

Kirke PN, Mills JL, Molloy AM, Brody LC, O'Leary VB, Daly L, Murray S, Conley M, Mayne PD, Smith O, Scott JM (2004): Impact of the MTHFR C677T polymorphism on risk of neural tube defects: case-control study. *BMJ*. 328:1535-6.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, Ferreira de Lima RLL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, et al (2002): Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. (Letter) *Nature Genet* 32:285-289.

Kruglyak L (1999): Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 22:139-44.

Kyvik KO, Green A, Beck-Nielsen H (1995): Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. *BMJ* 311:913-7.

Laitinen T (2002): The value of isolated populations in genetic studies of allergic diseases. *Curr Opin Allergy Clin Immunol* 2:379-382.

Lander ES, Schork NJ (1994): Genetic dissection of complex traits. *Science* 265:2037-2048.

Laumonnier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, Raynaud M, Ronce N, Lemonnier E, Calvas P, Laudier B, Chelly J, Fryns JP, Ropers HH, Hamel BC, Andres C, Barthelemy C, Moraine C, Briault S (2004): X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet* 74:552-7.

Leoyklang P, Siriwan P, Shotelersuk V (2006): A mutation of the p63 gene in non-syndromic cleft lip. (Letter) *J Med Genet* 43:e28.

Leroy JG (2007): Oligosaccharidoses, disorders allied to the oligosaccharides; Ch 108 in: Rimoin DL, Connor JM, Pyeritz RE, Korf BR (eds) *Emery and Rimoin's Principles and Practice of Medical Genetics*, 5 ed. Vol 3. Churchill Livingstone Elsevier, Philadelphia, pp 2413-48.

Lesage S, Zouali H, Cezard J-P & the EPWG-IBG group, Colombel, J-F & the EPIMAD group, Belaiche J & the GETAID group, Almer S, Tysk C, O'Morain C, Gassull M, Binder V *et al.* (2002): CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. *Am J Hum Genet* 70:845-857.

Losh M, Sullivan PF, Trembath D, Piven J (2008): Current developments in the genetics of autism: from phenome to genome. *J Neuropathol Exp Neurol* 67:829-37.

Mannion JJ (1977): *The peopling of Newfoundland: Essays in Historic Geography*, St. John's, Memorial University of Newfoundland.

Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Feibig A, Schreiber S, Friedman J, Ketelaars CEJ, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW (2008): Structural Variation of Chromosomes in Autism Spectrum Disorder. *Am J Hum Genet* 82:1-12.

Martin LJ, Crawford MH, Koertvelyessy T, Keeping D, Collins M, Huntsman R (2000): The population structure of ten Newfoundland outports. *Hum Biol* 72:997-1016.

McCandless SE, Brunger JW, Cassidy SB (2004): The burden of genetic disease on inpatient care in a children's hospital. *Am J Hum Genet* 74:121-7.

McDonald-McGinn DM, Emanuel BS, Zackai EH (2005): 22q11.2 Deletion syndrome. University of Washington. GeneTests. Available at <http://www.genetests.org/>

McKusick's Online Mendelian Inheritance in Man (OMIM™): <http://www.ncbi.nlm.nih.gov/Omim/>

Merner ND, Hodgkinson KA, Haywood AF, Connors S, French VM, Drenckhahn JD, Kupprion C, Ramadanova K, Thierfelder L, McKenna W, Gallagher B, Morris-Larkin L, Bassett AS, Parfrey PS, Young TL (2008): Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene. *Am J Hum Genet* 82:809-21.

Meunke M, Gropman A (2005): Holoprosencephaly overview. University of Washington. GeneTests. Available at <http://www.genetests.org/>

Miles J, McCathern (2005): Autism Overview. University of Washington. GeneTests. Available at <http://www.genetests.org/>

Miles JH, Takahashi TN, Bsgby S, Sahota PK, Vaslow DF, Wang CH, Hillman RE, Farmer JE (2005): Essential versus complex autism: definition of fundamental prognostic subtypes. *Am J Med Genet* 135A:171-80.

Moore SJ, Buckley DJ, MacMillan A, Marshall HD, Steele L, Ray PN, Nawaz Z, Baskin B, Frecker M, Carr SM, Ives E, Parfrey PS (2008): The clinical and genetic epidemiology of neuronal ceroid lipofuscinosis in Newfoundland. *Clin Genet* 74:213-22.

Moore SJ, Green JS, Fan Y, Bhogal AK, Dicks E, Fernandez BA, Stefanelli M, Murphy C, Cramer BC, Dean JCS, Beals PL, Katsanis N, Bassett AS, Davidson WS, Parfrey PS. (2005): The clinical and genetic epidemiology of Bardet-Biedel syndrome in Newfoundland: A 22-year prospective, population-based, cohort study. *Am J Med Genet* 132A: 352-360.

MRC (1991): Prevention of neural tube defects. *Lancet* 338:131-137.

Nall L, Gulliver W, Charmley P, Farber EM (1999): Search for the psoriasis susceptibility gene: the Newfoundland Study. *Cutis* 64:323-9.

Newhook LA, Curtis J, Hagerty D, Grant M, Patterson AD, Crummel C, Bridger T, Parfrey P (2004): High incidence of childhood type I diabetes in the Avalon Peninsula of Newfoundland, Canada. *Diabetes Care* 27:885-888.

Newhook LA, Grant M, Sloka S, Hoque M, Paterson AD, Hagerty D, Curtis J (2008): Very high and increasing incidence of type 1 diabetes mellitus in Newfoundland and Labrador, Canada. *Pediatric Diabetes* 9:62-8.

Noble, JA et al. (1996): The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. *Am J Hum Genet* 59: 1134-1148.

Nussbaum RL¹, McInnes RR, Willard HF (2001): Genetic Variation in Populations; Ch 7 in *Thompson and Thompson: Genetics in Medicine*, 6 ed, W.B. Saunders Company, pp. 105-107.

Nussbaum RL², McInnes RR, Willard HF (2001): Principles of Clinical Cytogenetics; Ch 9 in *Thompson and Thompson: Genetics in Medicine*, 6 ed, W.B. Saunders Company, pp. 146-147.

Nussbaum RL³, McInnes RR, Willard HF (2001): Patterns of Single-Gene Inheritance; Ch 5 in *Thompson and Thompson: Genetics in Medicine*, 6 ed, W.B. Saunders Company, p 61-71.

Nussbaum RL⁴, McInnes RR, Willard HF (2001): Patterns of Single-Gene Inheritance; Ch 5 in *Thompson and Thompson: Genetics in Medicine*, 6 ed, W.B. Saunders Company, p 56-57.

Nussbaum RL⁵, McInnes RR, Willard HF (2001): Principles of Clinical Cytogenetics; Ch 9 in *Thompson and Thompson: Genetics in Medicine*, 6 ed, W.B. Saunders Company, pp. 150-157.

O'Dea D, Parfrey PS, Harnett JD, Hefferton D, Cramer BC, Green J (1996): The importance of renal impairment in the natural history of Bardet-Biedl syndrome. *Am J Kid Dis* 27:776-783.

Oen K, Robinson DB, Nickerson P, Katz SJ, Cheang M, Peschken CA, Canvin JM, Hitchon CA, Schroeder ML, El-Gabalawy HS (2005): Familial seropositive rheumatoid arthritis in North American Native families: effects of shared epitope and cytokine genotypes. *J Rheumatol* 32:983-91

Olney R, Mulinare J (2002): Trends in neural tube defect prevalence, folic acid supplementation, and vitamin supplement use. *Semin Perinatol* 26:277-285.

Olufemi SE, Green JS, Manickam P, Guru SC, Agarwal SK, Kester MB, Dong Q, Burns AL, Spiegel AM, Marx SJ, Collins FS, Chandrasekharappa SC (1998): Common ancestral mutation in the MEN1 gene is likely responsible for the prolactinoma variant of MEN1 (MEN1Burin) in four kindreds from Newfoundland. *Hum Mutat* 11:264-269.

Ou CY, Stevenson RE, Brown VK, Schwartz CE, Allen WP, Khoury MJ, Rozen R, Oakley GP Jr, Adams MJ Jr (1996): 5,10 Methylene tetrahydrofolate reductase genetic polymorphism as a risk factor for neural tube defects. *Am J Med Genet* 63:610-4.

Peltonen L, Jalanko A, Varilo T (1999): Molecular genetics of the Finnish disease heritage. *Hum Mol Genet.* 8:1913-23.

Persad VL, Van den Hof MC, Dube JM, Zimmer P (2002): Incidence of open neural tube defects in Nova Scotia after folic acid fortification. *Can Med Assoc J* 167: 241-245.

Persico AM, Burgeron T (2006): Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends Neurosci* 29:349-58

Poorthuis BJHM, Wevers R, Kleijer WJ, Groener JEM, de Jong JGN, van Weely S, Niezen-Koning KE, Van Diggelen OP (1999): The frequency of lysosomal diseases in The Netherlands. *Hum Genet* 105:151-6

Rahman P², Bartlett S, Siannis F, Pellet FJ, Farewell VT, Schentag CT, Alderdice CA, Hamilton S, Khraishi M, Tobin Y, Hefferton D, Gladmann DD (2003): CARD15 – a pleiotropic autoimmune gene that confers susceptibility to psoriatic arthritis. *Am J Hum Genet* 73:677-81.

Rahman P¹, Jones A, Curtis J, Bartlett S, Peddle L, Fernandez BA, Freimer NB (2003): The Newfoundland population: a unique resource for genetic investigation of complex diseases. *Hum Molec Genet* 12:R167-R172.

Reddy KS (2005): Cytogenetic abnormalities and fragile-X syndrome in Autism Spectrum Disorder. *BMC Med Genet* 6:3.

Risch N (1990): Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46:222-228.

Ross AJ, Beales PL (2007): GeneReviews Bardet-Biedl syndrome. Available at <http://www.genetests.org/>

Schaefer GB, Mendelsohn NJ (2008): Genetics evaluation for the etiologic diagnosis of autism spectrum disorders. *Genetics in Medicine* 10:4-12.

Scriber CA, Neal JL, Saginur R, and Clow A (1973): The frequency of genetic disease and congenital malformation among patients in a pediatric hospital. *CMA Journal* 108: 1111-1115.

Shaffer LG, Bejjani BA (2006): Medical applications of array CGH and the transformation of clinical cytogenetics. *Cytogenet Genome Res.* 115:303-9.

Sheehy MJ, Scharf SJ, Rowe JR, Neme de Gimenez, MH, Meske LM, Erlich HA, Nepom BS (1989): A diabetes-susceptible HLA haplotype is best defined by a combination of HLA-DR and -DQ alleles. *J Clin Invest* 83:830-835.

Sieberts SK, Schadt EE (2007): Moving toward a system genetics view of disease. *Mamm Genome* 18:389-401.

Stankiewicz P, Beaudet AL (2007): Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation. *Curr Opin Genet Dev* 17:182-92.

Stankiewicz P, Lupski JR (2003): Genome architecture, rearrangements and genomic disorders. *Trends Genet.* 18:74-82.

Stevenson DA, Carey JC (2004): Contribution of malformations and genetic disorders to mortality in a children's hospital. *Am J Med Genet* 126A:393-7.

Stewart DL, Hersh JH (1995): The impact of major congenital malformations on mortality in a neonatal intensive care unit. *J Ky Med Assoc* 93:329-32.

Stuckless S, Parfrey PS, Woods MO, Cox J, Fitzgerald GW, Green JS, Green RC (2007): The phenotypic expression of three MSH2 mutations in large Newfoundland families with Lynch syndrome. *Fam Cancer* 6:1-12.

Tsui NB (2008): Multiple thiopurine S-methyltransferase variation detection: a step toward personalized medicine. *Clin Chem* 54:1637-47.

Varilo T, Peltonen L (2004): Isolates and their potential use in complex gene mapping efforts. *Curr Opin Genet Dev* 14:316-323.

Vermeire S, Wild G, Kocher K, Cousineau J, Dufresne L, Bitton A, Langelier D, Pare P, Lapointe G, Cohen A, Daly MJ, Rioux JD (2002): CARD15 genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, haplotype structure. *Am J Hum Genet* 71:74-83.

De Wals P, Rusen ID, Lee NS, Morin P, Niyonsenga T (2003): Trend in prevalence of neural tube defects in Quebec. *Birth Defects Res (Part A)* 67: 919-923.

De Wals P, Tairou F, Van Allen MI, Uh SH, Lowry RB, Sibbald B, Evans JA, Van den Hof MC, Zimmer P, Crowley M, Fernandez B, Lee NS, Niyonsenga T (2007): Reduction in neural-tube defects after folic acid fortification in Canada. *NEJM* 357: 135-42.

Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ; the Autism Consortium (2008): Association between Microdeletion and Microduplication at 16p11.2 and Autism. *N Engl J Med* 358:667-75.

Wilson RD, Johnson JA, Wyatt P, Allen V, Gagnon A, Langlois S, Blight C, Audibert F, Désilets V, Brock JA, Koren G, Goh YI, Nguyen P, Kapur B; Genetics Committee of the Society of Obstetricians and Gynaecologists of Canada and The Motherisk Program (2007): Pre-conceptional vitamin/folic acid supplementation 2007: the use of folic acid in combination with a multivitamin supplement for the prevention of neural tube defects and other congenital anomalies. *J Obstet Gynaecol Can* 29:1003-26.

Woodcock J (2007): The prospects for "personalized medicine" in drug development and drug therapy. *Clin Pharmacol Ther* 81:164-9.

Woods MO, Young TL, Parfrey PS, Hefferton D, Green JS, Davidson WS (1999): Genetic heterogeneity of Bardet-Biedl syndrome in a distinct Canadian population: evidence for a fifth locus. *Genomics* 55:2-9.

Xie YG, Zheng H, Leggo J, Scully MF, Lillicrap D (2002): A founder factor VIII mutation, valine 2016 to alanine, in a population with an extraordinarily high prevalence of mild hemophilia A. *Thromb Haemost* 87:178-9.

Yoon PW, Olney RS, Khoury MJ, Sappenfield WM, Chavez GF, Taylor D (1997): Contribution of birth defects and genetic diseases to pediatric hospitalizations. A population-based study. *Arch Pediatr Adolesc Med* 151:1096-103.

Young TL, Penney L, Woods MO, Parfrey PS, Green JS, Hefferton D, Davidson WS (1999): A fifth locus for Bardet-Biedl syndrome maps to chromosome 2q31. *Am J Hum Genet* 64:900-904.

Young, TL, Woods MO, Parfrey PS *et al.* (1998): Canadian Bardet-Biedl syndrome family reduces the critical region of BBS3 (3p) and presents with a variable phenotype. *Am J Med Genet* 78:461-7.

Zahir F, Friedman JM (2007): The impact of array genomic hybridization on mental retardation research: a review of current technologies and their clinical utility. *Clin Genet* 72:271-87.

Zbuk KM, Stein JL, Eng (2006): PTEN hamartoma tumor syndrome (PHTS). GeneTests. Available at <http://www.genetests.org/>

Zhou S (2006): Clinical pharmacogenomics of thiopurine S-methyltransferase. *Curr Clin Pharmacol* 1:119-28.

Zielenski J, Corey M, Rozmahel R, Markiewicz D, Aznarez I, Casals T, Larriba S, Mercier B, Cutting GR, Krebsova A, Macek M Jr, Langfelder-Schwind E, Marshall BC, DeCelle-Germana J, Claustres M, Palacio A, Bal J, Nowakowska A, Ferec C, Estivill X, Durie P, Tsui LC (1999): Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet* 22:128-9.

Category IA: Chromosomal disorders

Down syndrome (trisomy 21)
Trisomy 13
Trisomy 18
Turner syndrome
Microdeletion syndromes¹
Uniparental disomy
47, XYY
Ring chromosome mosaicism

Category IB: Single-gene disorders

Arginosuccinate deficiency
Arthrogryposis
Autosomal dominant polycystic kidney disease (ADPKD)
Caffey disease
Canavan disease
Cardiomyopathy (familial)
Cataracts (autosomal dominant)
Citrullinemia
Cleidocranial dysplasia
Congenital adrenal hyperplasia (CAH)
Cornelia de Lange syndrome
Craniotelsoncephalic dysplasia
Cutis laxa
Cystic fibrosis
Donahue syndrome
Ectodermal dysplasia
Factor V Leiden
Familial adenomatous polyposis syndrome (FAP)
Fanconi anemia
Glanzmann syndrome
Gordon syndrome
Hereditary angioedema
Hereditary hearing loss/deafness²
Hereditary spherocytosis
Hereditary retinoblastoma
Homocysteinuria

¹ E.g. Angelman/Prader Willi syndrome, 22q11.2 microdeletion, Williams syndrome.

² With mutant gene identified or with Mendelian family history.

Category IB: Single-gene disorders (continued)

I-cell disease (mucopolidosis type II)
Ivemark syndrome (AR)
Krabbes leukodystrophy
Lebers hereditary optic neuropathy
Lesch-Nyhan syndrome
Marfan syndrome
Medium chain acyl CoA dehydrogenase (MCAD) deficiency
Metachromatic leukodystrophy (MLD)
Multiple endocrine neoplasia type 1 (MEN1)
Mitochondrial myopathy
Nemaline myopathy
Neurofibromatosis type 1 (NF1)
Neuronal ceroid lipofuscinosis (infantile)
Noonan syndrome
Opitz syndrome
Osteogenic imperfecta
Osteopetrosis
Pelizaeus-Merzbacher disease
Peutz-Jeghers syndrome
Protein C deficiency
Pyruvate kinase deficiency
Retinitis pigmentosa
Rett syndrome
Romano-Ward syndrome
Saethre-Chotzen syndrome
Severe combined immunodeficiency syndrome (SCIDS)
Severe myoclonic epilepsy of infancy (SMEI)
Sotos syndrome
Tuberous sclerosis
Tyrosinemia

Category IC: Teratogens

Accutane embryopathy
Diabetic embryopathy
Fetal alcohol spectrum disorder (FASD)

Category ID: Genetic syndromes without identified gene and/or genetic mechanism

Goldenhar Syndrome (Hemifacial Microsomia)
Kabuki syndrome
VACTERL association

Category IE: Probable unidentified genetic syndrome

Examples:

Child with unexplained developmental delay and dysmorphic facial features.

Child with unexplained developmental delay and birth defects.

Congenital myopathy (no muscle biopsy).

Category IIA: Multifactorial birth defect with known recurrence risk to siblings (as per Harper 2001):

Branchial cleft sinus
Cleft lip/palate
Clubfoot
Congenital cataract
Congenital glaucoma
Congenital heart disease
Congenital hip dysplasia
Craniosynostosis
Cryptorchidism
Dandy-Walker malformation syndrome
Ear tags/pits
Encephalocele
Hearing loss³
Hemihypertrophy
Hirschsprung disease
Holoprosencephaly
Hypospadias
Microcephaly
Lissencephaly
Neural tube defect
Ossicular fusion
Prune belly syndrome
Pyloric stenosis
Renal agenesis
Situs inversus

³ Congenital hearing loss not meeting the IB criteria and without obvious environmental cause (e.g. prematurity).

Category IIB: Birth defect with low recurrence risk to siblings
(as per Harper, 2001)

Agensis of corpus callosum
Arachnoid cyst
Biliary atresia
Bladder extrophy
Bowel atresia
Bowel malrotation
Cerebral dysgenesis
Cervical thymic cyst
Congenital diaphragmatic hernia
Congenital nevus
Diastomyelia
Gastroschisis
Horseshoe Kidney
Hydrocoele
Hydronephrosis
Imperforate anus
Inguinal hernia
Leg length discrepancy
Miscellaneous external ear malformations (including ear sinus)
Multicystic dysplastic kidney
Omphalocele
Pilonidal sinus
Posterior urethral valves
Radial ray defects
Sacral agenesis
Strabismus
Subarachnoid cyst
Syndactyly
Syrinx of spinal cord
Tarsal coalition
Tracheoesophageal fistula (TEF)
Urachal remnant

Category IIIA: Diseases with known genetic predisposition
[including with heritability (h^2) > 50% or lambda s > 10]

Alcohol dependence
Ankylosing spondylitis
Anorexia nervosa
Asthma
Attention deficit disorder/ Attention deficit hyperactivity disorder
Autism spectrum disorder
Cataract- not congenital
Celiac disease
Cervical and lumbar disc degeneration
Conduct disorder
Depression (childhood onset)
Diabetes- Type 1
Graves disease
Hypothyroidism (including congenital)
Inflammatory bowel disease (Crohns disease or ulcerative colitis)
Juvenile rheumatoid arthritis
Mental retardation/developmental delay
Migraine
Multiple sclerosis
Obesity
Schizophrenia
Scoliosis
Seizure disorders (febrile or afebrile)⁴
Systemic lupus erythematosus (SLE)
Tourette syndrome
Vesicoureteric reflux (includes hydronephrosis due to reflux)

⁴ Excludes first seizure which is classified as IIIB.

Category IIIB: Diseases with multiple known causes, sometimes genetic
[including those with heritability (h^2) < 50% or $\lambda_s < 10$]

Allergies (environmental)
Anemia
Anomalies of dental arch
Angioedema (sporadic)
Anxiety disorders (panic disorder)
Arrhythmias
Atopic dermatitis (eczema)
Blindness
Cancer
Cerebrovascular accident (CVA) - includes cerebral infarct
Cholesteatoma
Constipation
Dental caries
Diabetes – type II
Diabetes gestational
Disseminated intravascular coagulation (DIC)
Encopresis
Eosinophilic cystitis
Failure to thrive (FTT)
Gastroesophageal reflux
Glaucoma (not congenital)
Hearing loss⁵
Horners syndrome
Hypertension, essential
Hypertension, pregnancy induced
Hypopituitarism
Hypotonia
Idiopathic thrombocytopenia purpura (ITP)
IgA deficiency
Lactose intolerance
Legge-Perthes disease
Motor delay
Neurogenic bladder
Osteoarthritis
Patent ductus arteriosus (not requiring surgery)
Peptic ulcer disease
Prematurity
Renal stones
Rhabdomyolysis
Seizure (first)

⁵ Hearing loss not meeting the IB or IIA criteria.

Serum sickness

Wolf-Parkinson-White syndrome

Category IV – Acquired disease with low or no genetic contribution

Burns

Trauma

Infection (e.g. pneumonia, infectious gastroenteritis)

Chemotherapy-related complications

Category V – No disease

Appendix B – ANCOVA
Case Mixed Group (CMG) Codes

Table A-1: Numbers of patients from unique admission dataset of 3,281 patients belonging to each of the 20 most common Case Mixed Group (CMG) codes.

20 most common CMG's (code number)	Number of unique patients
Seizures & headache (22)	99
Mastoid procedures (85)	41
External & middle ear procedures (90)	35
Myringotomy (92)	44
Tonsillectomy & adenoidectomy procedures (93)	374
Influenza (104)	40
Simple Pneumonia & pleurisy (143)	152
Tracheobronchitis (145)	148
Asthma (146)	140
Simple Appendectomy (262)	38
Esophagitis, gastroenteritis & Misc. digestive disease (294)	311
Cellulitis (447)	32
Diabetes (483)	54
Major urinary tract procedures (504)	27
Lower urinary tract infection (529)	33
Upper extremity procedure for trauma (670)	37
Other cranial injuries (695)	40
Viral illness (757)	37
Fever of unknown origin (761)	40
Disruptive behavior disorders (786)	37
Other CMG codes	1,522
Total	3,281

Day and Holmes 1973 - Reanalysis of the frequency of genetic disease for 200 pediatric admissions using the present study's genetic content categories

Chromosomal disorders (our category IA) – no admissions

Mendelian diseases (our category IB) – 8 admissions (4%)

Environmental (our category IV) – 92 admissions (46%) – no diagnoses specified

Table C-1: *Number and diagnoses of patients from Day and Holmes 1973 study belonging to our genetic content categories IIA, IIB, IIIA and IIIB*

Current study category IIA: (Diagnosis – number of patients)	Current study category IIB: (Diagnosis – number of patients)	Current study category IIIA: (Diagnosis – number of patients)	Current study category IIIB: (Diagnosis – number of patients)
NTD – 1 Encephalocoele – 1 Congenital heart disease - 1 Pyloric stenosis – 3 Cong. dislocated hip – 1 Cleft palate – 1 Cryptorchism -1	Urinary tract abnormality – 36 Thyroglossal duct cyst – 1 Hemangioma – 1 Cystic hygroma – 2 Aneurismal bone cyst – 1 Agenesis left lung – 1 Ectopic anus – 1	Asthma – 4 Psoriasis – 1 Diabetes type I – 2 Seizures – 1 Migraine – 3 Parry-Rhomberg syndrome – 1 Congenital encephalopathy – 1	Hernia – 9 Cerebral palsy – 3 Wilms tumor – 3 Peripheral vasomotor instability – 1 Chondromalacia – 1 Chronic renal disease - 1
Total 22 (11%)	Total 43 (21.5%)	Total 1 (8.5%)	Total 18 (9%)

Hall et al. 1978 - Reanalysis of the frequency of genetic disease for 4,115 admissions using the present study's genetic content categories

Chromosomal (our study category IA) = 26 patients (0.6%)

Mendelian (our study category IB) = 162 patients (3.9%)

Teratogen (our study category IC) = 11 patients (0.27%)

Multiple anomalies suggestive of a genetic syndrome (IE) = 10 patients (0.24%)

There were no diagnoses identified from Hall's diagnoses lists which would have been classified in the present study as ID (syndromes of unknown genetic etiology).

IIA (high heritability birth defects) = 421 (10.2%), as shown in table A-3 below:

Table C-2: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIA (high heritability birth defects).

Hall's category	Diagnosis	Frequency (number of patients)
IIA – multifactorial established	Cleft lip +/- palate	51
	Congenital heart malformation (no specific diagnoses given)	162
	Congenital dislocated hip	47
	Meningomyelocele	44
	Club feet	20
	Hirschsprung disease	15
	Pyloric stenosis	14
IIB – presumed multifactorial	Hydrocephalus	18
III –developmental anomalies	Cryptorchidism	25
	Hypospadias	25
Total belonging to present study category IIA		421

Table C-3: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIB (low heritability birth defects).

Hall's category	Diagnosis	Frequency (number of patients)
III –developmental anomalies	Genitourinary anomalies (renal)	48
	Genitourinary anomalies (other genital excluding cryptorchidism & hypospadias)	18
	Other genitourinary anomalies	43
	Eye (including esotropia)	65
	Gastrointestinal anomalies (including TE fistula and imperforate anus)	41
	Skeletal (including limb abnormalities)	41
	Head and neck (including ear anomalies)	40
	Skin (including hemangiomas)	33
	CNS anomalies	15
	Miscellaneous	37
Total belonging to present study category IIB		381 (9.3% of her dataset)

Table C-4: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIIA (high heritability multifactorial diseases excluding birth defects).

Hall's category	Diagnosis	Frequency (number of patients)
IIA – multifactorial established	Allergy (including asthma)	157
	Diabetes type I	50
IIb –presumed multifactorial	Seizures	137
	Scoliosis	26
	Mental retardation	24
	Others including migraine, RA, ulcerative colitis	28
IV - familial	Collagen vascular disorders	19
Total belonging to present study category IIIA		441 (10.7% of her dataset)

Table C-5: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IIIB (low heritability multifactorial diseases excluding birth defects).

Hall's category	Diagnosis	Frequency (number of patients)
IIA – multifactorial established	Duodenal ulcer	4
	Hypertension	2
IIb –presumed multifactorial	Cerebral palsy	99
III –developmental anomalies	Hernias	118
IV – familial	Cancers and tumors	230
	Renal disorders	108
	Prematurity	45
	Orthopedic disorders	41
	Hematologic	33
	RDS	12
	Newborn jaundice	10
	Aborted SIDS	8
	Miscellaneous	37
Total belonging to present study category IIIB		747 (18.2%)

Table C-6: Number and diagnoses of patients from Hall et al. 1978 belonging to our genetic content category IV (acquired disease with low or no genetic contribution).

Hall's category	Diagnosis	Frequency (number of patients)
V – non-genetic	Infectious diseases	1,194
	Trauma	349
	Ingestions	80
	GI symptoms	68
	Burns	48
	Deficiency states	36
	Appendicitis	32
	Iatrogenic or post-operative complications	29
	Miscellaneous	51
Total belonging to present study category IV		1,887 (45.9% of her dataset)

Healthy child (our genetic content category V) = 29 patients (0.7%)

Table C-7: Summary of the number of patients from Hall et al. 1978 belonging to our 11 genetic content categories.

Present study Category	Number of patients in Hall et al. dataset	Frequency (%)
IA	26	0.63
IB	162	3.9
IC	11	0.27
ID	-	-
IE	10	0.24
IIA	421	10.2
IIB	381	9.3
IIIA	441	10.7
IIIB	747	18.2
IV	1887	45.9
V	29	0.70
TOTAL	4,115	~100

Appendix C-3:
Comparison of Categories from McCandless et al. 2004 to present study

1. McCandless category IA = chromosomal and single gene disorders; 622 cases or 10.8% of their dataset. This corresponds to categories IA, IB, ID and IE of present study (i.e. 8.4% of our consecutive separation dataset)
2. McCandless category IB (*"multifactorial / polygenic"*): 823 cases or 14.5% of their dataset. All the McCandless IB diagnoses are included in our category II A (high heritability birth defect), except the following which we categorized as IIIA: depression; seizures; celiac disease; scoliosis.
3. McCandless category IC (*"heterogeneous causes often genetic"*): 495 cases or 8.6% of their dataset. We would have categorized these individual diagnoses as a mixture of categories IIA, IIB, IIIA, IIIB, ID and IE as shown below in *table C-8*:

Table C-8: Diagnoses of patients from McCandless category IC belonging to our genetic content categories IIA, IIB, IIIA, IIIB, ID and IE.

McCandless IC diagnoses classified as IIA in present study	McCandless IC diagnoses classified as IIB in present study	McCandless IC diagnoses classified as IIIA in present study	McCandless IC diagnoses classified as IIIB in present study	McCandless IC diagnoses classified as ID in present study	McCandless IC diagnoses classified as IE in present study
Congenital nystagmus; Congenital cataracts; Craniosynostosis; Hearing loss*; Microcephaly;	Atresia large bowel; Arteriovenous malformation; Bicornuate uterus; Bone cysts; Branchial cleft cyst; Congenital heart block; Congenital hydrometrocypous; Congenital diaphragmatic hernia; Hydronephrosis; Laryngotracheomalacia Mandible hypoplasia; Microcolon; Pectus excavatum; Moya-Moya; Multicystic dysplastic kidney; Radioulnar synostosis Umbilical hernia; Velopharyngeal insufficiency	Adolescent kyphosis; Alopecia; Aortic aneurysm; Autism; Developmental Delay; Hypothyroidism; Migraine; Mental retardation; Obesity; Recurrent joint dislocation	ALTE ; Endometriosis ; ESRD; FTT; Membranoprolif. Glomerulonephritis; Myoglobinuria; Nasal sinus polyp; Osteochondroma; Ovarian cysts; Pes plannus; Recurrent pneumothorax; Recurrent vomiting; Retinal detachment; Supraventricular tachycardia; Wilms tumor	Oculoauriculo Vertebral spectrum; VACTERL association	Arthrogryposis

* We classified hearing loss as IB if the child had a Mendelian family history or had a mutation in one of the deafness genes. Hearing loss was classified IIA if it was congenital without obvious environmental cause and did not meet the IB criteria. Otherwise hearing loss was categorized as IIIB.

Appendix C-3:
Comparison of Categories from McCandless et al. 2004 to present study

4. McCandless category IIA (*"malformations of unknown etiology"*) contained 48 cases or 0.84% of their total separations. It partially overlapped with our category IIB (low heritability birth defects) as shown below in *table C-9*.

Table C-9 *Diagnoses of patients from McCandless category IIA belonging to our genetic content categories IIB, IIA and ID.*

McCandless IIA diagnoses classified as IIB in present study (i.e. agreement)	McCandless IIA diagnoses classified as IIA in present study (i.e. disagreement)	McCandless IIA diagnoses classified as ID in present study (i.e. disagreement)
Bladder extrophy Hydrocoele Meckel diverticulum	Branchial cleft sinus Chiari I malformation	Sturge-Weber syndrome

5. McCandless category IIB (*"teratogens"*) completely agreed with our category IC. Their data set contained 19 cases (0.33% of their total hospitalizations). Our dataset contained 8 cases or 0.2 % of our total hospitalizations.
6. McCandless category III (*acquired disorder with genetic predisposition*) contained 2,096 cases (36.5% of total cases). It was a mixture of our categories IIIA and IIIB (higher and lower heritability multifactorial diseases, excluding birth defects) as shown below in *table C-10*.

Table C-10: *Diagnoses of patients from McCandless category III belonging to our genetic content categories IIIA and IIIB.*

McCandless III diagnoses classified as IIIA in present study	McCandless III diagnoses classified as IIIB in present study
Asthma; Crohns disease; Graves disease; Inflammatory bowel disease; Type I diabetes; Juvenile rheumatoid arthritis; Developmental delay of unknown cause; Schizophrenia; Depression; Anorexia nervosa; Milk protein intolerance; SLE; Suicide; Ulcerative colitis	ALL; Benign bone tumor; CML; Exostosis; Germ cell tumor; Hodgkins lymphoma; Malignancy; Neuroblastoma



